MANANGING ALGORITHMIC METRICS AND CUSTOMERS: A MULTI-CASE STUDY OF LABOR CONTROL AND RESISTANCE IN THE GIG ECONOMY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ngai Keung Chan

May 2021

MANANGING ALGORITHMIC METRICS AND CUSTOMERS: A MULTI-CASE STUDY OF LABOR CONTROL AND RESISTANCE IN THE GIG ECONOMY

Ngai Keung Chan, Ph. D.

Cornell University 2021

This dissertation examines how algorithmic metrics become sources of labor control and resistance in the gig economy. Using a qualitative multi-case study approach, this dissertation investigates (a) how labor platforms use metrics to govern their distributed workforces and (b) how digitally-enabled service workers encounter, interpret, and manage metrics as part of their daily work practices and service interactions. Drawing on the discursive analysis of platforms' corporate discourses and in-depth interviews with 50 workers in the United States, this project examines and compares workers' practices across three kinds of labor platforms, namely, ride-hailing platforms, TaskRabbit, and delivery platforms. These platforms are built around algorithmic metrics that entail constant monitoring and surveillance of workers' performance, while platform-based labor processes and the visibility of metrics vary across the three cases. I argue that metrics, particularly customer ratings, rationalize and reinforce the management of workers by customers through the production of work-related uncertainty and anxiety. Meanwhile, I find that workers learn to strategically manage their customers before, during, and after service interactions to reduce work-related uncertainty, and thus maintain their autonomy. I

also discuss how worker-to-worker online communities emerge as important social spaces for workers to share strategies and feel connected. Furthermore, I explore how socio-technical features of labor platforms and workers' economic dependence on the platform shape the disciplinary outcomes of metrics. Taken together, this dissertation offers a comparative lens for understanding the role of algorithmic metrics in shaping the trilateral relationship between platform owners, service workers, and customers in the gig economy. It also underscores the need to rethink how metrics, and more broadly, digital data transform the service triangle in workplaces.

BIOGRAPHICAL SKETCH

Ngai Keung Chan conducted his PhD study in the Department of Communication at Cornell University. Prior to Cornell, he completed both his undergraduate degree in Government and Public Administration and MPhil in Communication at The Chinese University of Hong Kong.

This dissertation is dedicated to Bazla Kadir for her incredible love and support.

ACKNOWLEDGMENTS

This dissertation is a collaborative effort. I have been incredibly fortunate to have a terrific dissertation committee: Lee Humphreys, Brooke Erin Duffy, Malte Ziewitz, and Tarleton Gillespie. I express my deepest gratitude to Lee, my brilliant advisor and mentor, for her mentorship and endless intellectual and emotional support. I still remember the first time we met (virtually) during the admission interview in January 2016. This was a major reason I decided to come to Cornell. And this is one of the best life decisions I have ever made. Lee has mentored me with wisdom and enormous patience. Lee, thank you for being an amazing role model to me! I am also immensely grateful to Brooke for guiding me through different stages of intellectual and professional development. Brooke, it was my privilege to work with you in various capacities—in class, working as your teaching and research assistant, and collaborating with you—and I learned so much about what it means to be a scholar. Malte, thank you for being my constant source of support and inspiration! Your course on evaluation and society was one of my favorite classes at Cornell and has really opened my eyes to the field of evaluation studies and ignited my research interest. Tarleton, thank you for gently pushing me to think more deeply about my theoretical and analytical approaches (aka "Gillespie-ish" framework). Although we worked at different locations, you were always very generous with your time and insights. I hope I can be the mentor to others that all of you have been to me someday.

I am also thankful to the Department of Communication—a warm and caring community. I am particularly grateful to Jeff Niederdeppe. Jeff, thank you for giving me the opportunity to teach your course on communication campaigns. I learned a lot

from your teaching and advice. I would also like to thank the New Media and Society Working Group—including Katherine Sender, Chelsea Butkowski, Megan Sawey, Colten Meisner, Sara Eddleman, Beatrys Rodrigues, Lucas Wright, Leah Scolere, Elizabeth Newbury, and Sandy Payette (and thanks again to Lee and Brooke!)—for providing unfailing support, sharing their inspirational work, and offering thoughtful feedback on my earlier drafts. I also want to thank all of the administrative staff in the Department of Communication, specifically Joanna Alario who was always there to help. The graduate student community was another wonderful source of support for me. I would like to thank Wen Duan, Luping Wang, Motasem Kalaji, Salma El idrissi, Julie Cannon, Amanda Purington, Negar Khojasteh, Samuel Hardman Taylor, Ge Gao, Byungdoo Kim, Vincent Yu, Swati Pandita, Melissa Seipel, and Xiaoyan Li.

I would also like to thank my friends and collaborators who are always understanding, encouraging, and supportive. Special thanks to Chi Kwok, Chris Chao Su, Alice Renwen Zhang, and Krishnan Vasudevan.

Finally, I want to thank for my family for their continued support. My eternal gratitude goes to my partner and soulmate Bazla Kadir. I cannot imagine having completed this dissertation without her humor, love, sacrifices, and care.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Noel[1] was an Uber driver in Louisville and completed over 4,400 rides between 2016 and 2018. He started driving for Uber because the platform allowed him to work on his own schedule and spend more time with his family. He considered Uber as a part-time job and set a goal to make $100 per day. He maintained an almost perfect customer rating—4.98 out of 5-star rating for his last 500 rated ride—but he still paid close attention to his rating. The rating, for him, was "a reflection of the service that you provide and a reflection of your reputation as a driver amongst all the other drivers." He told me, "it will ruin my whole day if I get a bad rating." Despite his high rating, he said, "[t]here's always room for improvement each time … That's my goal to improve my rating as a driver, because that's my reputation. That's what motivates me to be a good driver. It's my rating." To this end, he stated, "I try to treat my customers with the most respect that I possibly can … I treat them as if they were my boss."

Carter was a DoorDash delivery driver in Massachusetts. He began working on the platform in 2018. His mother and younger brother were also DoorDash drivers. Alongside his full-time job as a medical biller, he worked for 40 hours on DoorDash in a typical week. While he enjoyed interacting with the customers, he was under pressure to maintain the overall customer rating. As he described, the customer rating was "a very fuzzy number." He said, "A couple of months ago I was shunted from 4.60 to 4.75. And within a day … two days later, my customer rating dropped by 4.60,

[1] I assigned pseudonyms to all of the interviewees to protect their identities.

which I couldn't even explain it." The platform did not explain the sudden change in his rating and rarely provided customer feedback. The lack of such information, for Carter, was "the biggest problem" on DoorDash and prevented him for "improving" himself "to do better job next time."

Unlike working for Uber or DoorDash, Kooper and other Taskers can set their hourly rate and communicate with their prospective clients before accepting a request on TaskRabbit. While Uber and DoorDash rely on ratings to determine whether one can continue working on the platform, TaskRabbit's metrics indirectly affect the platform's suggested hourly rate for Taskers and the visibility of their online profiles. Compared to Noel's and Carter's fierce attention to their ratings, Kooper—a nine-year Tasker who was mainly hired for house cleaning, furniture assembly, moving, and delivery services in Los Angeles—seemed to care less about his metrics. Many of the metrics, such as the acceptance rate and response rate, were out of his control. Specifically, TaskRabbit's algorithm determines how quickly a Tasker responds to requests on the platform, whereas customer feedback only affects the Tasker's reliability rate and 5-stars rating. Yet, metrics and service interactions remained important in Kooper's experience. As he said, "If you have over a certain rating, your profile is higher in the search engines." When asked about his strategies for getting positive ratings, he said, "over-communication has been my trademark." He worked hard to manage the expectations of his customers before and during service interactions, including discussions about the price threshold, the length of the task, and the details of the task.

Noel, Carter, and Kooper worked for different on-demand labor platforms

where their platform-based metrics, and more broadly, digital data are subject to distinct design and incentive structures (i.e., rewards and sanctions). Nevertheless, they had similar struggles to interpret and manage their metrics in their daily work. In this dissertation, I use "metrics" as an umbrella term to describe the quantitative outputs about one's performance that are generated and processed by labor platforms. Platforms also determine the (in)visibility of metrics; specifically, what metrics make visible, to whom, and for what purposes. Here, metrics refer to socio-technical artifacts that classify, commensurate, and discipline workers' differences at a distance through the production and distribution of numbers. In the gig economy, the most common metrics are customer-sourced ratings—such as the one discussed in the stories from Noel and Carter—which enable service recipients to evaluate workers' performance, often using a common, numeral 5-stars scale. Yet, metrics do not necessarily involve a direct input from customers; for example, platforms may rely on their algorithms to calculate one's acceptance rate, cancellation rate, and response rate. As such, all of these metrics I studied in this project are algorithmic, as they are the outputs generated by algorithms and made visible to workers. In this project, I argue that metrics are not simply about "subjective" or "objective" representation of one's performance. Instead, I focus on processes through which metrics re-shape power-relations between platforms, workers, and customers, especially when metrics are connected with algorithmic direction and discipline.

Quantification and evaluation have long been central to labor management in manufacturing and service workplaces, closely associated with processes of surveillance discipline (Ajunwa et al., 2017; Ball, 2010; Braverman, 1974; Edwards,

1984; Frenkel et al., 1995; Moore, 2018). The emergence of the gig economy amplifies labor control through the reliance on algorithmic technologies (Gandini, 2019; Kellogg et al., 2020; Vallas & Schor, 2020). On the one hand, labor platforms deploy algorithmic metrics to monitor their distributed workforce (Gandini, 2019; Rosenblat, 2018a; Rosenblat & Stark, 2016; Ticona & Mateescu, 2018). On the other hand, workers respond to such technologies of evaluation through compliance, informal collaboration, gaming, and resistance, at individual and collective levels (Anwar & Graham, 2020; Bucher et al., 2021; Cant, 2020; Chen, 2018; Gray & Suri, 2019; Petre et al., 2019; Rosenblat, 2018a; Scholz, 2017; Ziewitz, 2019).

Against this backdrop, this dissertation explores the tensions between metrics as technologies of labor control and workers' reactive practices towards metrics across labor platforms in the gig economy in the United States. This project focuses on what I call *digitally-enabled service workers*. Regardless of working as an Uber driver, a DoorDash delivery driver, or a TaskRabbit Tasker, their primary tasks are matched and mediated via digital labor platforms (De Stefano, 2016) and involve direct social interactions with customers. Workers and customers encounter digital data *before* interacting with one another. Hence, digital data, particularly metrics may set their expectations of social interactions.

There are three reasons for studying metrics and digitally-enabled service workers in the gig economy. First, theorizations of the "reputation society" (Masum & Tovey, 2011), "reputation economy" (Gandini, 2016; Hearn, 2010), "metric society" (Mau, 2019) and "metric power" (Beer, 2016) highlight how metrics become emergent forms of capital that re-organize our social and economic life (see also Fourcade &

Healy, 2017). Metrics predate the digital era. Taylorism and scientific management represent a precursor of the quantification of work. That said, not all forms of metrics are generated by algorithms. Productivity metrics, for instance, can be simply expressed as the numbers of units a worker produce within a certain time frame. This does not necessarily involve any computational calculation and evaluation.

Outside of the workplace, credit ratings, censuses, and bestseller lists for cultural products[2] have existed for a while (Blank, 2007; Carruthers, 2013; Hearn, 2010; Mennicken & Espeland, 2019). Dating back to the 1990s, e-commerce sites such as eBay developed online ratings and reviews to build trust among buyers and sellers (Dellarocas, 2003). In anonymous online marketplaces, buyers and sellers are total strangers. If there is a lack of trust, economic transactions will not happen. From an economic perspective, online ratings and reviews—which can be considered "digitalized word-of-mouth (Dellarocas, 2003)—can hold buyers and sellers accountable by making their reputational information visible to a community of users (Tadelis, 2016). Online rating and reviews can also generate implicit and explicit signals to help socialize new users to conform to the community standard (Lampe, 2012). More recently, we are witnessing a rapid expansion of digital metrics outside of e-commerce sites (Esposito & Stark, 2019). Popular media narratives—such as the *Black Mirror*'s "Nosedive" and coverages about the gig economy (e.g., Dzieza, 2015; Mahdawi, 2016)—have envisioned dystopian futuristic worlds wherein our life chances are dependent on how others rate us. An inquiry into gig workers' practices

[2] Consider, for instance, audience measurement ratings which can be used to measure and represent audiences' media exposure to a particular cultural product (Buzzard, 2012).

contributes to broader scholarly and popular debates on the impacts of metrification.

Second, the workplace is one of the spheres that is increasingly datafied: algorithms and the digital data they process affect labor processes from hiring (Ajunwa & Greene, 2019; Raghavan et al., 2020) to work assignment (Gandini, 2019; Lee et al., 2015) to performance evaluation and job termination (Kellogg et al., 2020; Rosenblat, 2018a). The deployment of algorithmic technologies does not only exist in expert occupations but also service workplaces. Yet, compared to knowledge workers such as educators (Espeland & Sauder, 2007, 2016; Muller, 2018), medical professionals (Maiers, 2017), and journalists (Anderson, 2011; Christin, 2018, 2020a; Christin & Petre, 2020), service workers have less autonomy (Ticona, 2015) to resist the influences of algorithmic metrics, due to the lack of recourse, organizational support, and professional expertise.

Third, the emphasis on *service workers* allows us to situate metrics into a "three-way contest for control and satisfaction" (Leidner, 1993, p. 22) in service labor processes. As sociologists of service work have long argued, employers, workers, and customers have their own interests, which can affect the performance of emotional and interactive labor (Hochschild, 1983/2012; Korczynski, 2009a; Leidner, 1993; Lopez, 2010). The routinization of service interactions requires the platform, workers, and customers to coordinate their performance. In theory, metrics not only enable labor management *by* customers but also incite the management *of* customers. As in the stories of Noel and Kooper, they learned to manage their customers and service interactions, in part due to the pressure for scoring well. Metrics, meanwhile, are what Beer (2016) calls "affective measures": "[m]etrics are communication with purse and

intention" that provokes particular forms of emotional responses to favor productivity and market competition (p. 197; see also Espeland & Sauder, 2016; Mau, 2019; Muller, 2018). The feelings of uncertainty and anxiety incite workers to play by the rules, strategically manage their customers, and resist the influences of metrics. Although many labor platforms use metrics to transform workers' performance into numbers, metrics do not necessarily come with same disciplinary outcomes. Metrics matter when workers are aware of being measured and take them seriously as part of their daily work practices.

The primary objective of this project, therefore, is to provide a comparative assessment of the power dynamics underpinning the management, interpretation, and uses of metrics in platform-specific calculative spaces. An over-arching question that guides this project is: How do algorithmic metrics perform as technologies of control and resistance in gig workers' everyday work practices? Here, resistance is broadly defined as workers' attempts to maintain work autonomy and re-claim a sense of control over the labor process. An underlying premise is that the powerless groups, particularly service workers often undertake hidden, offstage, and subtle forms of resistance to protect themselves from sanctions (Anwar & Graham, 2020; Hodson, 1995; Scott, 1990; Ticona, 2015; Woodcock, 2017). This conceptualization enables me to examine different moments where workers attempt to claim their agency (Woodcock, 2017). Based on this over-arching question, I explore the following sub-questions:

1. How do platforms use metrics to govern their distributed workforces?

2. How do workers interpret and manage metrics as part of their daily service interactions and work practices?
3. What are the factors that shape workers' interpretation and uses of metrics across labor platforms?

To address these questions, I use a multi-case study approach to explore gig workers' practices around three types of labor platforms: ride-hailing platforms (Uber and Lyft); TaskRabbit, an on-demand temporary (temp) work platform; and on-demand delivery platforms (DoorDash, Uber Eats, and Instacart). Using in-depth interviews with 50 gig workers and the discursive texts related to these platforms (e.g., company documents and popular press articles), I examine and compare (a) how platforms justify their uses of algorithmic metrics and (b) how workers strategically interpret and use these metrics for practical purposes. While existing studies have examined the tensions between platform governance and workers' practices on labor platforms discretely, this project offers a comparative assessment of workers practices across labor platforms. Indeed, a few studies (Gray & Suri, 2019; Ravenelle, 2019; Schor, 2020; Wood et al., 2019) conduct comparative analysis of labor platforms and offer valuable insights into the impacts of algorithms on workers. Building on these studies, I focus on how metrics, and more broadly, algorithmic evaluation takes place and shape labor processes in the gig economy. This project argues for the need to rethink how metrics, and more broadly, digital data transform the service triangle in workplaces. As many of the interviews with workers were conducted during the COVID-19 pandemic, this project may also contribute to understanding how workers navigate metrics amidst the time of crisis.

I will describe each case later in this chapter and offer a methodological justification in Chapter 3. In the three cases, workers must interact with digital data and their prospective customers in their daily work. Each case represents a distinct type of service work, incentive structures (i.e., rewards and sanctions that are connected with metrics), and visibility of metrics. For example, TaskRabbit Taskers' metrics are visible to their prospective clients, whereas Uber and Lyft drivers' metrics remain only visible to themselves in most cases. Moreover, metrics have various direct and indirect outcomes across these platforms. These differences can be usefully analyzed as distinct socio-technical assemblages (Ettlinger, 2018; Gillespie, 2014; Kitchin, 2017) around workers' interpretation and uses of metrics. All of the platforms I studied, except Instacart, offer rewards programs that require workers to maintain a certain level of metrics. The comparative analysis allows us to paint a vivid picture of why metrics matter and how they may reinforce unequal power-relations in the gig economy.

The Rise of the "Gig Economy"

I now turn to discuss what I mean by the gig economy and situate it into broader popular and scholarly debates around non-standard and precarious employment. Following Woodcock and Graham (2020)'s conceptualization, I define the gig economy as "labor markets that are characterized by independent contracting that happens through, via, and on digital platforms" (p. 11). There are various definitions of platforms, as corporations strategically frame and legitimize their services and data practices on a discursive level (Gillespie, 2010). According to van Dijck, Poell, and de Waal (2018), platforms are "fueled by *data*, automated and

organized through *algorithms and interfaces,* formalized through *ownership* relations driven by *business models,* and governed through *user agreements*" (p. 8; italics original). In the gig economy, labor platforms are firms that provides socio-technical infrastructures (e.g., mobile applications and algorithms). They rely on such infrastructures and user agreements for organizing labor processes. Meanwhile, platforms are the sites where workers encounter digital data and use them for work-related purposes. In other words, labor platforms' algorithms match and mediate economic transactions between independent contractors and service recipients (e.g., Uber/Lyft rides, Uber Eats/ DoorDash/ Instacart delivery orders, and TaskRabbit tasks) (Vallas & Schor, 2020; Woodcock & Graham, 2020). Under this conceptualization, work tasks can be performed offline or online, but this project mainly discusses gig work that involves some forms of in-person service interactions with customers.[3] In what follows, I want to briefly discuss popular discourses about the gig economy to contextualize the project.

Over the past decade, the "gig economy" has emerged as a buzzword to highlight how the rise of digitally-enabled independent contracting and short-term work arrangements is "disrupting" our workplaces and potentially becoming "the future of work" (Horowitz, 2011; Mulcahy, 2019; Schwartz et al., 2016; Volkin, 2020). "Gig"—originally a slang that means "the short-term arrangements typical of a musical event" (Woodcock & Graham, 2020, p. 11; see also Baym, 2018)—has been

[3] Examples of gig work that is performed entirely online include freelancing platforms (e.g., Upwork) and on-demand crowdwork or what Gray and Suri (2019) call ghost work platforms (e.g., Amazon Mechanical Turk). Indeed, the typology of crowdwork is also in flux; for instance, Howcroft and Bergvall-Kåreborn (2019) consider Uber and TaskRabbit as a form of asset-based crowdwork.

rebranded as flexible employment opportunities and entrepreneurship (Prassl, 2018) in popular business discourses: be your own boss and work anytime you want. For example, Sara Horowitz, the founder of Freelancers Union, wrote in 2011,

> It's been called the Gig Economy, Freelance Nation, the Rise of the Creative Class, and the e-conomy, with the "e" standing for electronic, entrepreneurial, or perhaps eclectic. … No longer do we work at the same company for 25 years, waiting for the gold watch, expecting the benefits and security that come with full-time employment. We're no longer simply lawyers, or photographers, or writers. Instead, we're part-time lawyers-cum-amateur photographers who write on the side … This transition is nothing less than a revolution.

In 2012, Diane Mulcahy, an adjunct lecturer at Babson College and a management consultant, began to teach a course entitled "Entrepreneurship and the Gig Economy" (Cheng, 2019; Mulcahy, 2019) and later published a book to advise people on how to thrive in the gig economy. Cheng (2019), a reporter for *Quartz*, found that college career services centers at the University of Texas-Austin (UT Austin) and Wellesley College started to offer advice on gig work. Not surprisingly, the two quoted career services centers focused more on freelancing career than digitally-enabled service work. Yet, UT Austin's Liberal Arts Career Services suggested students "driving for ride-hailing companies or working for delivery services like Instacart," when they need a part-time job to "make extra money while in school" (Cheng, 2019).

Alongside popular business discourses, "Uber for X" (Schor, 2020; Srnicek, 2017) was once a prominent business model that attempted to pitch for venture capital funding. As Webb (2016) found on a database for startups and angel investors, "526

companies included “Uber for” in their listings,” varying from medical services to legal services to cleaning to dog walking. In 2019, Madrigal (2019) tracked the development of 105 “Uber for X” companies in the United States and found that they received $7.4 billion investment from venture capital firms. About half of these startups no longer exist or are acquired by other companies. For instance, TaskRabbit, one of the cases in this dissertation, was acquired by IKEA in 2017. This is not to argue for or against the “Uberization” discourse (Scholz, 2017; c.f. Ticona & Mateescu, 2018), but to highlight that all of these independent and temporary work—from freelancing to on-demand legal services to ride-hailing—fall under the umbrella of the gig economy.

The vague definition results in varied estimates about the size of the gig economy, varying from 0.7% to 40% of the American workforce. Other reasons for the difficulty in estimation are that gig workers may consider platform work as a side gig and they work on multiple platforms at the same time (Woodcock & Graham, 2020). The “Gig Economy Data Hub” (2020) collected a variety of data sources about the gig economy. The largest estimate (i.e., 40%) came from the analysis of the 2010 General Social Survey data, but it included part-time workers, contract company workers, and independent contractors (US Government Accountability Office, 2015). Drawing on a sample of 39 million Chase checking accounts between 2012 and 2018, the JP Morgan Chase Institute found that about by March 2018, 1.6% of adults had generated income from an online platform at least once in the past month (Farrell et al., 2018). The transportation platforms accounted for over half of the transactions (56%; about 150 million dollars). Situating the gig economy into the global context,

Oxford's "Online Labour Index" (Kässi & Lehdonvirta, 2018) tracks changes in the number of freelancing tasks and projects on Fiverr, Freelancer, Guru, and PeoplePerHour. Between May 2016 and January 2021, the number of new tasks has grown about 40%. Drawing on a review of existing studies of the gig economy in North America and Europe, an OECD study found that the digitally-enabled gig economy accounted for approximately 1% to 3% of employment (Schwellnus et al., 2019).[4] A more recent estimate from The Fairwork Project (2020) suggests that there are up to 55 million gig workers around the world.[5] The estimated size of the gig economy has likely gone up during the COVID-19 pandemic.

What do the varying estimates tell us about the gig economy? First, at least millions of gig workers are, to a certain extent, working on a digital labor platform where they must make sense of the platform-based algorithms and data. Second, the rough estimates suggest that the gig economy only accounts for a small portion of the workforce and it is not yet replacing traditional employment, though it seems to prevail in corporate discourses about the "future of work" (van Doorn, 2017). It is out of the scope of this dissertation to interrogate such discursive work (Prassl, 2018), but such discourses may speak to a wider concern over the expansion of precarious work (Ravenelle, 2019; Scholz, 2017; Vallas & Schor, 2020; Woodcock & Graham, 2020)

[4] The definition of the gig economy in this report is similar to the one that I use in this dissertation. Schwellnus et al. (2019) define gig economy platforms as "two-sided digital platforms that match workers on one side of the market to customers (final consumers or businesses) on the other side on a per-service ("gig") basis (p. 6).

[5] The estimate was ranged from 14 to 55 million gig workers. The Fairwork Project (2020) suggests "the upper number increases to nearly 440 million workers when country-specific estimations for the size of local gig economies are incorporated into the calculation" (p. 12). It claims that the details about the calculations will be provided in a forthcoming paper.

and potentially the fear of having an "algorithmic boss" (Rosenblat, 2018a).

What is New About the Gig Economy?

Short-term, on-demand, and temporary work arrangements are core characteristics of the gig economy. Platforms position themselves as digital intermediaries between the demand and supply of labor rather than employers to distance to shift risks and responsibilities to workers (Bieber & Moggia, 2020; Fleming, 2017; McDonald et al., 2020; Stanford, 2017; van Doorn, 2017). From the perspective of platform owners, labor platforms are "disruptive" innovations that promise to liberate independent contractors from traditional bureaucratic labor markets. DoorDash's (n.d.-a) sign-up page for drivers has the tagline: "Your time. Your goals. You're the boss." What a driver needs "is a mode of transportation and a smartphone to start making money." TaskRabbit (n.d.-a) conveys a similar message and states, "With TaskRabbit, you have the freedom and support to be your own boss." While much has been written about how platforms fail to realize such promises (e.g., Ravenelle, 2019; Rosenblat, 2018a), it is vital to note that people's desire for job flexibility is one of the motivating factors that make the gig economy possible (Woodcock & Graham, 2020). Throughout my research, the interviewees like Noel frequently mentioned flexibility as a key reason for working on the platform.

Scholars have argued that gig work, or more precisely, precarious work is neither novel nor innovative. Precarious work, explain Kalleberg and Vallas (2018, p. 1), refers to "work that is *uncertain, unstable, and insecure and in which employees bear the risks of work … and receive limited social benefits and statutory protections*" (italics original). It has been expanding across occupations since the 1970s, driven by

de-unionization, financialization (e.g., the emergence of outsourcing and downsizing), globalization, and digitalization (Kalleberg, 2009; Kalleberg & Vallas, 2018). Indeed, precarious work has an even longer history than standard employment. As Kalleberg (2009) argues, "the postwar period (up until the mid-1970s) was unusual for its sustained growth and stability" (p. 5; see also Neilson & Rossiter, 2008). Much like other forms of precarious workers (Standing, 2011), gig workers often lack labor-related protection such as employment, economic, and representation securities (Gregory, 2021; McDonald et al., 2020; Ravenelle, 2019). The gig economy, therefore, may simply reveal "a digital version of the offline atypical, casual, freelance, or contingent work arrangements" (Kalleberg & Dunn, 2016, p. 11).

What distinguishes gig work from other forms of precarious work is that labor platforms offer "a convenient, readily available infrastructure with which to limit the firm's obligation to the workforce on which it relies" (Vallas & Schor, 2020, p. 278; see also Kalleberg & Dunn, 2016). There are two notable aspects of platforms' infrastructures in the scholarly literature. First, platforms are "multi-sided" (Rochet & Tirole, 2003), meaning that their businesses are only valuable when they can attract a large number of workers and consumers to use their infrastructures. Pinpointing platforms as "a new type of firm" (p. 48), Srnicek (2017) argues that platforms rely on the accumulation of network effects and data extraction to achieve market dominance. As a monopolistic labor platform can reach a massive amount of consumers and workers, it has the potential to transform the economy sectors they enter and alter regulatory infrastructures in favor of their operation (Culpepper & Thelen, 2020;

Pollman & Barry, 2017; Rahman & Thelen, 2019).[6]

Second, of particular importance in this dissertation is concerned with the role of algorithms and digital data in transforming labor processes in the gig economy. Algorithms seemingly replace human managers to direct, evaluate, and discipline gig workers (Kellogg et al., 2020; Lee et al., 2015; Rosenblat, 2018a), which can in turn result in workers' feelings of instability and insecurity (Duffy, 2020; Malin & Chandler, 2017). Of course, this is not the first time "new" technologies have reorganized work activities (e.g., Barley, 1985; Hall, 2010; Smith, 2015). Corporations, for example, often introduce surveillance technologies for monitoring the pacing of work tasks in the name of care (Sewell & Barker, 2006). Long before the emergence of the gig economy, workers are reportedly subject to what Frenkel et al. (1995) call "info-normative control," meaning that the seemingly "objective" performance indicators set normative standards for controlling workers' performance.

Nonetheless, algorithmic management of digitally-enabled service workers rests upon a trilateral relationship between platform owners, workers, and customers. On the one hand, platform owners must consider how to measure and evaluate the performance of their distributed workforces. The evaluation of workers is not just about labor management, but also about building trust with consumers and maintaining the quality of services in anonymous marketplaces. On the other hand, it is true that platform owners can continuously collect workers' real-time data at a much

[6] As I focus on gig workers' daily work practices, I do not plan to discuss the politics of platform power here. Yet, I want to highlight this strand of literature as they point to a distinctive feature of platform firms, as compared to traditional consolidated firms (e.g., General Motors) (Rahman and Thelen, 2019).

larger scale than before (Kellogg et al., 2020). Such data, however, do not and cannot guarantee that all service interactions are under total control and surveillance. Therefore, platform owners need to enroll workers and customers "in the practices of algorithmic management without managerial authority having been delegated to them" (Stark & Pais, 2021, p. 47).

Algorithmic metrics, the objects under study, represent a key mechanism for platforms to create calculative spaces for workers to modify their behaviors in response to being evaluated by their customers. The information visibility of metrics varies across labor platforms (e.g., Ticona et al., 2018). It is against this backdrop that I examine the disciplinary impacts of algorithmic metrics. Like any forms of standardization and quantification, metrics are "imperfect systems." A standardized evaluation system must define what counts and what does not count to transform workers' performance in different service interactions into numerical values that are commensurable across social contexts (Espeland & Stevens, 2008; Esposito & Stark, 2019; Fourcade, 2011; Lamont, 2012). Nonetheless, gig workers must learn to navigate such "imperfect" systems and manage service interactions. How, then, do metrics become a form of social knowledge that enable labor platforms to exercise uneven influences on workers' practices? Accordingly, this dissertation considers the implications of algorithmic labor control and resistance in the gig economy.

Current Trends in the Gig Economy

Throughout my research between 2017 and 2021, there are three important trends in shaping the development of the gig economy. These on-going trends are essential for understanding the empirical importance of the gig economy and

contextualizing this project.

First, several labor platforms have become public companies via an initial public offering (IPO). Ari Levy (2020) created a list of the biggest IPOs for tech companies by 2020. Lyft went public in March 2019 and raised $2.6 billion (ranked no. 8 all-time by 2020). Uber is the second biggest tech IPO (since Facebook in 2012) and raised $8.1 billion in 2019. DoorDash's 2020 IPO also raised $3.4 billion (ranked no. 7). Instacart is also planning to go public in 2021 (Bosa & Batchelor, 2021). Going public does not necessarily mean that their businesses are profitable. In 2018, 81% of U.S. companies filing IPOs were unprofitable (Molla, 2019). The gig economy is no exception. Uber, for example, had a net loss of $1.1 billion in the first three quarters of 2020 (Uber Investor, 2020). Although DoorDash significantly expands its market share in on-demand delivery services due to the COVID-19 pandemic, its IPO prospectus revealed a net loss of $149 million in 2020 (Feiner, 2020). Yet, these companies' IPO prospectuses help us to better understand their business models: labor platforms' anticipated profitability is grounded in the worker classification of independent contractors (see also Prassl, 2018). For instance, Lyft's (2019) IPO prospectus claimed that if drivers are classified as employees, it "could harm our business, financial condition and results of operation" (p. 28). DoorDash (2020a) also spent at least $30 million to support the "Prop 22" campaign in California to "obtain certainty regarding worker classification matters" (p. 148).

This leads to the second trend about employment classification of gig workers. Although I focus on gig workers' day-to-day contestation of algorithmic labor management, there are efforts underway to push for the reclassification of gig

workers. One notable move is the passage of the California Assembly Bill (AB5) in September 2019. This bill shall require labor platforms to reclassify gig workers as employees, meaning that gig workers would have access to employment benefits such as minimum wage, paid sick leave, and unemployment insurance. As a response, Uber and Lyft threatened to shut down their services in California (Hawkins, 2020a). Later, Uber, Lyft, DoorDash, Instacart, and Postmates sponsored the Proposition 22 (Prop 22) campaign, a ballot that pushed back the AB5. The campaign spent over $205 million, making it the most expensive ballot measure in California since 1999. It was eventually passed in January 2021. While there are on-going lawsuits against the ballot (Siddiqui, 2021), these series of events reveal the emerging political power of labor platforms. Gig workers, however, remain classified as independent contractors.

Third, gig workers have been exposed to heightened health and financial risks during the COVID-19 pandemic (The Fairwork Project, 2020). This situation has become worse for ride-hailing drivers and gig work that requires a significant amount of face-to-face interactions (Marshall, 2020; Marshall & Baker, 2020). In March 2020, the Independent Drivers Guild's survey found that drivers' weekly income decreased 60% in New York City (Marshall, 2020). Furthermore, the University of Chicago Harris School of Public Policy and the Associated Press-NORC Center for Public Affairs Research (2020) conducted a nationwide survey and found that fewer Americans used ride-hailing services during the pandemic (from 42% before the pandemic to 16% after the pandemic). The demand for delivery services remained largely the same. However, young adults and people with higher incomes reported an increased frequency of using grocery delivery services. The supply of delivery

workers, meanwhile, is burgeoning: Uber Eats reportedly had 36,000 new drivers in New York between March and November 2020 (Freytas-Tamura, 2020), whereas Instacart had doubled its shoppers to 500,000 since the pandemic (Lavietes & McCoy, 2020). As I will discuss in Chapter 3, this trend may be reflected in the interview sample: many of the delivery workers I interviewed were newcomers, but they did not plan to stay long in the gig economy. Workers' dependence on the platforms, therefore, shaped their reactive practices towards metrics.

Studying Platforms in a Comparative Perspective: An Overview of the Cases

I study three types of digitally-enabled service work. Table 1 offers an overview of each case. For each case, I conducted in-depth interviews with gig workers and drew on company websites and popular press articles to explore the tensions between algorithmic labor control and resistance.

Table 1. Overview of Cases

Platforms	Service work	Algorithmic matching process	Visibility of metrics	Metrics-related rewards and sanctions for workers
Uber and Lyft	Ride-hailing	Platforms match a rider to a driver who is logged into the app.	Only visible to the account owners and customers during a ride	Both platforms have drivers rewards programs. Platforms may deactivate drivers with a rating lower than a minimum threshold.
TaskRabbit	Temporary work	Workers create their online profile, availability, and rate. Clients can select and hire Taskers for a task.	Visible to the account owners and all prospective clients	TaskRabbit has an "Elite Tasker" program for its "highest rated Taskers." Taskers' ratings may also influence the visibility of their online profiles on the platform.

DoorDash, Uber Eats, and Instacart	Food and grocery delivery	DoorDash and Uber Eats match a customer to a driver who is logged into the app. Instacart shoppers can select batches in real-time.	Only visible to the account owners	DoorDash and Uber Eats have delivery drivers rewards programs. Instacart does not have a rewards program, but shoppers' ratings affect what batches they can view. Platforms may deactivate workers with a rating lower than a minimum threshold.

Uber and Lyft, the two major ride-hailing platforms, represent the first type of digitally-enabled service work. From 2017 to 2020, I conducted in-depth interviews with 21 Uber and/or Lyft drivers. Uber and Lyft have a similar labor process: using their mobile apps, drivers can provide ride-hailing services to other users who request a ride. Uber (2019a) and Lyft (2019) claim that their data-driven matching algorithms could generate precise predications about the demand and supply of ride-hailing services in real-time. According to the estimate of Second Measure (Yeo, 2021), Uber had about 71% of ride-hailing spending in the United States, whereas Lyft had the remaining 29% by December 2020. Although Uber and Lyft are competitors, the majority of my interviewees (14 of 21) worked on the two platforms at the same time. The work tasks of Uber and Lyft drivers are similar to that of taxi drivers, but their performance is tracked by 5-star ratings, acceptance rates, and cancellation rates. Both platforms allow drivers and riders to rate each other's performance on a 5-star scale. Uber calculates a driver's average ratings based on the last 500 rated trips, whereas Lyft only accounts for an average of the last 100 ratings.

Despite the difference in Uber's and Lyft's methods of calculation, their driver

metrics have relatively a low level of visibility. Drivers can see their metrics, whereas riders can look at their drivers' 5-star ratings during their ride. Uber and Lyft use drivers' ratings to determine drivers' employment opportunities and information visibility. Additionally, Uber and Lyft have similar design in terms of their platform interfaces and rewards program. Lyft's rewards program has three tiers (Silver, Gold, and Platinum), while Uber Pro program has four tiers (Blue, Gold, Platinum, and Diamond). Lyft's rewards program requires drivers to have a minimum rating of 4.9, whereas Uber drivers must maintain at least a 4.85 rating to unlock their tier status. Alongside the minimum rating thresholds, drivers need to earn points to unlock tiers in a given 3-month period. Drivers can see more trip information before accepting a ride, as they unlock tiers. However, drivers' tiers are not directly connected with their pay on the platform. By contrast, when drivers have an average rating lower than the location-specific minimum threshold, they may receive a warning message from the platform or be deactivated.

Founded in 2008, TaskRabbit is an on-demand temporary work platform. I conducted interviews with 11 Taskers in 2020. As Scholz (2017) argues, "TaskRabbit is now much closer to the model of a regular temp agency" (p. 48; see also Ravenelle, 2019). Taskers can be hired for a variety of tasks such as moving, cleaning, and furniture assembly. Some of these tasks are outsourced by corporations. Unlike the other two cases, the work of Taskers involves a significant amount of communication with their clients *before* accepting and completing a Task. This is because clients can have different expectations about a task; for example, how long a task should take and what tools a Tasker should prepare for and use for a task. After completing a task,

Taskers will be rated by their clients on a 5-star scale. Taskers need to create their online profile, set their hourly pay rate, and post their availability on the platform. A Tasker's profile also indicates whether s/he is an "Elite Tasker" and shows individual reviews and percentage of "positive" ratings (3 to 5 stars out of a 5-star scale). Therefore, metrics have a high level of visibility to Taskers and their prospective clients on TaskRabbit. Additionally, the platform controls the visibility of Taskers' online profile. When customers look for a Tasker for certain tasks, the platform only shows a recommendation list of potential Taskers. Customers can also select if they only look for "Elite Taskers" who are active on the platform each month and have a high rating. In short, metrics have an indirect impact on the search result of Taskers' profiles.

The third type of platforms I examined is on-demand delivery services, including Uber Eats, DoorDash, and Instacart. I interviewed 18 delivery workers from these platforms and the majority of them worked for either DoorDash or Uber Eats (15 of 18). DoorDash and Uber Eats focuses primarily on on-demand food delivery services, whereas Instacart offers grocery delivery services. As Table 1 shows, Instacart's labor process is slightly different from DoorDash and Uber Eats, as Instacart shoppers' ratings could potentially determine what batches (i.e., delivery orders) they can view on the platform. DoorDash and Uber Eats drivers, by contrast, can only decide whether to accept or decline an order matched by the platforms. However, at the time when I began to recruit Instacart shoppers for this project, Instacart (2020a) had already stopped using ratings to determine the visibility of batches. Additionally, DoorDash and Instacart use a 5-star rating system, whereas

Uber Eats asks customers to give thumbs up or thumbs down to drivers. Yet, all of which have a low visibility of metrics; specifically, only delivery drivers can check their metrics. Similar to ride-hailing platforms, delivery workers' ratings directly affect their employment opportunities.

Overall, each of the three cases highlight a distinct type of digitally-enabled service work (i.e., ride-hailing, temp work, and delivery). Service interactions are essential for workers to maintain their ratings in the three cases, though the technological design and incentive structures vary across labor platforms. Generally, there are two types of metrics across these labor platforms. Customer ratings entail human assessment of service quality, whereas acceptance rates, cancellation rates, and other system-generated metrics often focus on behavioral measures. The former incites workers to manage their service interactions because they are only rated *after* accepting and performing a task. Customer ratings are often used to determine one's continuity of employment on the platform. Acceptance rates and cancellation rates, meanwhile, may still impact workers' practices when considering their autonomy of accepting a task, in part because platforms do not make clear how metrics may affect the process of task allocation (Veen et al., 2020) and in part because they can affect workers' qualification for the rewards programs.

The ways that platforms use these two types of metrics are somewhat "black boxed" (Pasquale, 2015), meaning that the inner workings of metrics are opaque. For example, customer ratings and system-generated metrics are two of the factors that influence the visibility of Taskers' online profiles, but there are many other factors in TaskRabbit's matching algorithm. Furthermore, platforms can change the visibility of

metrics and modify their deactivation policy, without consulting workers. Workers must tap the in-app button to "agree" on the terms and conditions before working on the platforms. As platforms change the design and policies about metrics, workers need to re-acquire the knowledge about metrics and adapt their practices for practical purposes. This comparative project, therefore, examines the conditions and processes through which metrics become salient in shaping the power dynamics of platform owners, workers, and customers in the contexts of digitally-enabled service work.

Dissertation Outline

This dissertation consists of seven chapters. The introductory chapter and Chapter 2 contextualize the empirical site—digitally-enabled labor platforms in the gig economy—and introduce the main concepts. The third chapter reflects on my methodological approach. Chapters 4 to 6 constitute the core arguments of the dissertation: (a) why metrics are, in practice, "imperfect" systems that exercise labor control and rationalize the power of customers; and (b) how gig workers navigate metrics through the management of customers and information-sharing practices. The empirical chapters are organized by themes rather than cases. As I discussed earlier, there are differences in terms of the types of service work and the socio-technical features of labor platforms. However, the findings suggest that the interviewees experienced similar struggle to manage metrics and their customers. The final chapter concludes by reflecting on the broader implications of this project.

Chapter 2 reviews the key strands of literature that guided this project, including the power and reactivity of metrics, quantification and performance metrics in the workplace, service triangle in service work, and algorithmic labor control and

resistance in the gig economy. I also define the concepts of labor control and resistance that will be used throughout the dissertation. Additionally, I discuss the connections between the literature and my research project.

In Chapter 3, I discuss my methodological approach that guided the data collection and analysis of this project. I present the rationale for using a qualitative multi-case study approach. Then, I document the process of gaining access and data collection as well as the challenges that I encountered during the research process. I also describe the interview sample. After discussing my approach to data analysis, I reflect on the challenges and successes of the research process.

Chapter 4, "Algorithmic Labor Management and Metrics in Context," provides a brief history of the labor platforms I studied. Drawing on the analysis of discursive texts related to the labor platforms, I explicate the similarities and differences across the cases, focusing on the role of algorithmic metrics in the labor processes. Metrics as a form of algorithmic evaluation, I argue, is connected with algorithmic direction and discipline. Lastly, I discuss how these platform owners have legitimized their use of metrics, with upbeat assurances of mutual care and accountability.

Chapter 5, "The Power of Metrics and Customers," focuses on how metrics can create work-related uncertainty and feelings of anxiety across the three cases.[7] The empirical findings highlight how metrics rationalize and reinforce unequal power-

[7] Part of this chapter (i.e., the findings about ride-hailing drivers) was adapted from my published articles:

- Chan, N. K. (2019). The rating game: The discipline of Uber's user-generated ratings. *Surveillance & Society, 17*(1/2), 183–190. https://doi.org/10.24908/ss.v17i1/2.12911
- Chan, N. K., & Humphreys, L. (2018). Mediatization of social space and the case of Uber drivers. *Media and Communication, 6*(2), 29–38. http://dx.doi.org/10.17645/mac.v6i2.1316

relations between workers and customers. The analysis focuses on the affective power of customer ratings across the three cases. Nonetheless, not all metrics matter to the interviewees. I explore the factors that shaped the varying disciplinary impacts of metrics across the three cases.

While Chapter 5 reveals how algorithmic metrics enable the management *by* customers, I discuss how workers strategically manage their customers to navigate the "imperfect" metrics in Chapter 6.[8] I find that workers manage their customers *before*, *during*, and *after* service interactions. The management *of* customers allow gig workers to reduce work-related uncertainty around metrics. Additionally, particularly in the case of ride-hailing platforms, I document the emergence of worker-to-worker community as a way to empower workers to rework the labor process. I consider these practices as moments of resistance for workers to maintain autonomy in the gig economy.

Finally, in Chapter 7, I reflect on how metrics can be considered as technologies of control and resistance. After summarizing the findings, I discuss how this project contributes to understanding the trilateral relationship between platform owners, gig workers, and customers. I argue that metrics premediate gig workers' expectations of service interactions and incite them to engage with algorithmic labor control. Additionally, I discuss the limitations of this project. Moving forward, I suggest future research areas and connect them with the current trends in the gig

[8] Part of this chapter (i.e., the findings about ride-hailing drivers' strategies and Uber driver/bloggers' expertise) was adapted from my published articles including Chan (2019) and Chan, N. K. (2019). "Becoming an expert in driving for Uber": Uber driver/bloggers' performance of expertise and self-presentation on YouTube. *New Media & Society, 21*(9), 2048–2067. https://doi.org/10.1177/1461444819837736

economy. Importantly, this project considers how the disciplinary power of metrics can be extended to other forms of precarious service workplaces.

CHAPTER 2

LITERATURE REVIEW

The central question of this dissertation asks how algorithmic metrics constitute emerging forms of labor control and resistance in the gig economy. Addressing this question requires a clear conceptualization of metrics, labor control, and resistance. As I discussed in the first chapter, metrics are essentially about processes of quantifying workers' performance. It is noteworthy that not all metrics are algorithmic outputs, and vice versa. As in the history of quantification in the workplace, metrics such as quantified assessments of workers' timed performance have long existed before the digital era. Besides, algorithmic outputs can be the pricing of work tasks (e.g., rides or delivery orders) based on a number of factors such predictive demand for tasks. Such outputs are unrelated to workers' performance. In the gig economy, workers' metrics are either system-generated (e.g., acceptance rates) or customer-sourced (e.g., 5-star customer ratings), but all of them are processed by platform-based algorithms. Therefore, when I refer to metrics in this project, they are the outputs of platform-based algorithms. What distinguishes algorithmic metrics from previous forms of performance metrics is that the former is often used to automate employment decisions, which limit the ability for workers to appeal such disciplinary decisions.

I argue that the deployment of metrics has the potential to enable platform owners to set rules and norms to organize workers' activities and concentrate the ownership of job knowledge (i.e., labor control) (Edwards, 1984; Smith, 2015), while at the same time leaving room for workers to cope with the demands of their labor

(i.e., resistance) (Hodson, 1995; Scott, 1990). Therefore, the workplace is a "contested terrain" (Edwards, 1984) where "employers continuously innovate to maximize value captured from workers and workers inevitably engage in resistance to staining their autonomy, dignity, and identity" (Kellogg et al., 2020, p. 369).

In what follows, I review the key strands of literature that informed this project. I discuss why metrics matter in our social life, and specifically, in the workplace. Insofar as this project focuses on service workers, the concept of service triangle is useful for unpacking the trilateral relationship between employers, workers, and customers. Then, I explicate existing studies on algorithmic labor control and resistance in the gig economy. Lastly, I discuss how the literature informed my research questions.

The Power and Reactivity of Metrics

Our daily activities have been increasingly structured around ratings, rankings, likes, scores, and followers, among other quantified indicators (Esposito & Stark, 2019; Mennicken & Espeland, 2019; Mau, 2019; Muller, 2018). Metrics can seemingly measure, track, and predict one's creditworthiness (Carruthers, 2013; Kear, 2017), online influence (Gerlitz & Lury, 2014), reputation (Hearn, 2010; Masum & Tovey, 2011), and so forth. Metrics are inherently connected with the process of quantification—"the production and communication of numbers" (Espeland & Stevens, 2008, p. 8). Quantification transforms objects with different qualities into a numerical and categorical relationship, making them commensurable (Espeland & Stevens, 1998, 2008; Fourcade, 2011), by "conceal[ing] certain truths in revealing others" (Strathern, 2000, p. 315). In recent decades, we are witnessing an explosive

growth of metrics, metrics *of* metrics (e.g., the ranking of MBA rankings), and rating organizations in recent decades (Esposito & Stark, 2019).

There are at least four possible reasons for this trend. First, metrics, and more broadly, quantification has been developed in the context of political distrust, aspirating to replace personal and subjective expert judgment with standardized numbers (Porter, 1995; Muller, 2018). Metrics shall hold the authority accountable by providing "objective" information and "mak[ing] visible people, objects, or characteristics that may formerly have been invisible" (Espeland & Stevens, 2008, p. 415). Second, corporations can progressively use digital technologies to score and rank their customers for market-making purposes (Fourcade & Healy, 2017). Third, user-generated metrics are often connected with neoliberal discourses of consumer choice and empowerment (Kuehn, 2013; Mellet et al., 2014; Muller, 2018). Considered as "digitalized word-of-mouth" (Dellarocas, 2003), user-generated metrics and reviews have the potential to offer "perfect information" that informs consumers' decisions. Fourth, against the backdrop of information overflow, metrics can serve as "a point of orientation from which to navigate an otherwise uncertain decision space" (Esposito & Stark, 2019, p. 13).

Accordingly, metrics are supposed to be objective, transparent, accountable, and reliable, at least on a *discursive* level. It is also against this backdrop that metrics promise to re-create social order and inform decision-making processes. The ideal of mechanical objectivity legitimizes the uses of metrics for making "rational" decisions. As Porter (1995) puts this, "a decision made by the numbers (or by explicit rules of some other sort) has at least the appearance of being fair and impersonal" (p. 8). Once

the meanings of metrics are stabilized, they become a form of "mobile collectives" (O'Connell, 1993) that can be circulated and translated into various social contexts. Nonetheless, metrics are often "criticized for being simplistic, obscurantist, inaccurate, and subjective" (Esposito & Stark, 2019, p. 3). For example, people may carry implicit social biases when evaluating others (Rosenblat et al., 2017; Tadelis, 2016). Metrics as information signals, meanwhile, can guide people's expectations about the future, and therefore, their economic actions (Beckert, 2013; Salganik et al., 2006). Additionally, metrics define counts and wipe off alternatives. In the case of research assessment exercise, quantifiable indicators (e.g., publication scores) are devised to measure research outputs of academic institutions, but the indicators may eventually silence the social organization of institutions (e.g., the interactions between researchers) (Strathern, 2000). As such, metrics *inevitably* simplify complex situations for the sake of standardization and commensuration. As metrics are about the construction of anomalies, Star and Lampland (2009) have directed our attention to "what is being standardized, for what purpose, and with what result" (p. 15).

The power of metrics is more salient when metrics become a dominant mode of evaluation. Consider, for example, the plurality of rating and ranking systems. These evaluative systems may have different standards and compete against one another (Pollock et al., 2018). They also have unequal influences on organizations and individuals. In a study of organizational response to environmental ratings, Sharkey and Bromley (2015) found that firms would be more likely to change their behavior in response to ratings that "will prompt a reaction from consumers, competitors and regulators" (p. 87). By contrast, when a particular measure becomes a dominant

model, it can set normative standards for those who are being evaluated (Hoskin, 1996; Fourcade & Healy, 2017; Miller, 2001). Specifically, metrics become intertwined with responsibility, which prescribes what organizations and individuals *ought* to perform to improve their standings. The rankings of business schools (Wedlin, 2007) and law schools (Espeland & Sauder, 2007, 2016; Sauder & Espeland, 2009) exemplify this process. The rankings codify a "generalized" template for establishing the definition of "high-quality" education and promote intense competitions among schools.

What a performance measure intends to measure, in practice, is often not what is actually measured, in part because of the unintended interactions between the metrics and the objects being measured. This phenomenon has been explained through the lens of "reactivity," "Campbell's Law," and "Goodhart's Law." As Muller (2018) explains, "anything that can be measured and rewarded will be gamed" (p. 20). Metrics are "reactive," meaning that "people alter their behavior in reaction to being evaluated, observed, or measured" (Espeland & Sauder, 2007, p. 6). In their study of law school rankings, Espeland and Sauder (2016) found that the introduction of the rankings transformed the cognitive map that represents the hierarchical relationships (a) between law schools; (b) between law schools and the rankings; and (c) between law schools and external audiences. Law schools accordingly engaged in "reverse engineering," a recursive process of deconstructing the calculation of the ranking system and evaluating how resources could be effectively allocated to optimize their standing. The disciplinary power is not just about coercive pressures from universities and the public, but also deeply affective (Espeland & Sauder, 2016; see also Beer,

2016). Regardless of their rankings, law schools' reactive practices are driven by "the fear of falling in rank" (p. 4) and pressures of social comparison. Yet, the concept of reactivity suggests that organizations and people have the reflective agency to adjust their actions to comply with and optimize the perceived expectations of metrics, based on their understandings of what is being ranked as *relevant* and which relevant aspects are amenable in their situations.

Within the context of this project, it is important to consider how metrics can create disciplinary impacts on workers through the production of uncertainty and anxiety, on the one hand. On the other hand, the concept of reactivity directs our attention to how workers interpret a variety of metrics on labor platforms and strategically adapt their practices to manage metrics.

Quantification and Performance Metrics in the Workplace

The quantification of workers' performance has a long history. A well-known endeavor is the movement of "scientific management" initiated by Frederick Winslow Taylor in the late 19th and early 20th century. For Taylor (1911), "the best management is a true science, resting upon clearly defined laws, rules, and principles, as a foundation" (p. 7). This scientific management approach marked a departure from what Taylor called the management of "initiative and incentive" which had heavily rested upon giving special incentives and personal consideration for obtaining the "initiative" of workers (p. 34). Under scientific management, standardization and uniformity are the key to ensure workers to increase their productivity. To quote Taylor,

It is only through *enforced* standardization of methods, *enforced* adoption of

> the best implements and working conditions, and *enforced* cooperation that this faster work can be assured. And the duty of enforcing the adoption of standards and of enforcing this cooperation rests with the *management* alone. The *management* must supply continually one or more teachers to show each new man the new and simpler motions, and the slower men must be constantly watched and helped until they have risen to their proper speed. (p. 83; italics original)

To put it simply, the management has the responsibility to determine the most efficient way for distributing and standardizing work tasks. This involves the uses of a piece-rate system, time and motion studies, division of labor, and scientific selection and training of workers. As Braverman (1974) aptly argues, scientific management enables employers to "control each step of the labor process and its mode of execution" (p. 82) through the collection and abstraction of actionable knowledge about the labor process as well as the concentration of the knowledge in the hands of employers. In contrast to craft production, production is divided into step-by-step standardized, low-skill tasks. In this respect, quantification undermines the autonomy of front-line workers (Braverman, 1974; see also K. Levy, 2015; Moore, 2018).

Against the backdrop of the movement of scientific management, job evaluation was developed to determine wages in the United States in the 1920s and 1930s and became widely adopted by large manufacturers in the 1940s (Figart, 2001). The survey data from the National Industrial Conference Board showed that 75% of the sampled companies with more than 5,000 employees had job evaluation in 1946 (Baron et al., 1986). Figart (2001) documented that ranking, grade description, factor

comparison, and point factor were the four most common evaluation methods in the 1920s. The methods of ranking and grade description were rendered more "subjective" and less "scientific" as "[d]ecision makers considered their impressions of overall job content in hierarchically ranking or grading jobs" (Figart, 2001, p. 408). Indeed, the emergence of the other two methods, and more broadly, job evaluation was rhetorically framed as "equal pay for equal work" regardless of gender, race, and other characteristics of employees (p. 417). In practice, however, job evaluation instead reinforced gender pay inequalities (Figart, 2001; Steinberg, 1992) and helped corporations to lower the bargaining power of labor unions (Figart, 2001). Job evaluation carries employers' assumptions about what counts as "valuable" job skills (Steinberg, 1992). As such, job evaluation is inseparable from the continued control of workers.

Following Edwards (1984), a system of labor control includes three elements: directing work tasks, evaluating workers' performance, and disciplining workers. Edwards identifies three systems of labor control, namely, simple control (i.e., direct and arbitrary power), technical control, and bureaucratic control. The latter two types of control are particularly relevant to the quantification of work. Technical control refers to machinery pacing and organization of work tasks, whereas bureaucratic control entails the institutionalization of rules and hierarchical power-relations in workplaces. Although Edwards focuses on the systems of labor control in the 19th and 20th century, these two forms of control remain common in many service workplaces. Gottfried (1991), for instance, found that employers used a point system to allocate work tasks in the temporary help service industry. Points became intertwined with

what employers classified as “good conduct”; specifically, workers could earn points if they continuously work and accept every work assignment. The points could also be translated in retail value for purchasing items from the selected department stores. As such, the point system would reward those who frequently accepted tasks. Employers, meanwhile, relied on computers to monitor the pace of work. Likewise, managers rely on checkout operators to track the pacing of work tasks in an Australian supermarket chain (Price, 2016). Retail chains, such as Macy, also uses scorecards to monitor workers’ behavior and incorporate such metrics as part of workers’ annual performance reviews (Ikeler, 2016).

In contrast to Braverman’s (1974) emphasis on the degradation of work, Edwards (1984) contends that labor control and resistance are “two-way movement.” As quantification de-skills workers, workers can also re-skill themselves to resist the systems of labor control. Moreover, labor process involves both conflicts and coordination (Hyman, 1987). On the one hand, employers constantly reduce labor costs and optimize the labor process to maximize their earnings. On the other hand, an effective production cycle cannot operate without the coordination of workers. Therefore, labor control is not executed simply through coercion but an endeavor to win workers’ consent (Burawoy, 1979; Mears, 2015). This suggests the importance of considering how gig workers can maintain a sense of autonomy in the digitally-enabled labor process.

While the quantification of work is hardly new, the current trend of quantification places a greater emphasis on automated and algorithmic technologies (Moore, 2018). As Moore describes, “we are symbolically asked to “serve” the

machine" (p. 59). In the context of this project, gig workers, in most cases, are no longer monitored by human supervisors in the same location, but the algorithmic manager at a distance (Curchod et al., 2020; Lee et al., 2015; Rosenblat, 2018a). Algorithms *automatically* collect and analyze workers' performance per unit time and consumer-sourced feedback (Levy & Barocas, 2018). In doing so, it creates a gap of information visibility between platform owners, workers, and consumers (Beer, 2016; Curchod et al., 2020; Rosenblat & Stark, 2016). As Beer (2016) explains, "If there is one characteristic of metric power that we might want to take as central … it should be that of visibility and invisibility, attention and non-attention, recognition and non-recognition. It is such divisions that metric power gains its leverage" (p. 183).

In communication studies, scholars have directed attention to how workers interact with and manage metrics, particularly in newsrooms (e.g., Anderson, 2011; Christin, 2018, 2020a; Christin & Petre, 2020). Journalists are under organizational pressure to adopt quantitative data about audiences (e.g., ratings and the number of clicks and shares) in work settings. The click-based evaluation is often considered a threat to journalists' professional value and editorial autonomy. Journalists, therefore, develop relational strategies to resolve these tensions, for example, by drawing moral boundaries between "good" and "bad" metrics (Christin & Petre, 2020). Other studies examine how journalists' practice vary across organizational (Anderson, 2011) and national contexts (Christin, 2018, 2020a). Drawing on his ethnographic study of two newsrooms in Philadelphia and Newmark, Anderson (2011) found that journalists were less likely to rely on audience metrics to make news decisions, when there was a strong organizational culture against the dependence on metrics. Christin (2018,

2020a) demonstrates how national culture and the division of labor within newsrooms may affect journalists' reactive practices.

While these studies offer valuable insights into the interplay between metrics and worker autonomy, there is a major distinction between journalists and the gig workers I interviewed for this project. Journalists are considered "expert occupations" (Christin, 2017) with a relatively high level of professional autonomy and organizational support. Gig workers, by contrast, often work individually and lack organizational resources to adapt to the influences of metrics. This may suggest a different terrain of control and resistance in the gig economy. As on-demand labor platforms are built around performance metrics that enable algorithmic labor management (e.g., Gandini, 2019; Stark & Levy, 2018), they are valuable avenues for understanding the power and resistance of metrics in workplaces.

Service Triangle in Service Work

As this dissertation focuses on digitally-enabled service work, the concept of service triangle is helpful for considering power dynamics underlying the interactions between employers (i.e., platform owners), workers, and customers. Service work, explains Wharton (2015), is "inseparable from the person who produced them," meaning that "production and consumption can occur almost simultaneously in service encounters between workers and customers (or clients)" (p. 331). In other words, customers become part of service labor processes.

Leidner (1993), accordingly, introduces the concept of service triangle to understand the three-way interest alliance. To put this simply, employers, workers, and customers have their own interests. These interests can sometimes align with or

against one another, so they can affect the performance of emotional labor and service encounters (Leidner, 1993; Hochschild, 1983/2012). In order to routinize interactions, employers must consider how to script the role and actions of both workers and customers. As Leidner (1993) explains,

> In interactive service work, as in other kinds of work, understanding employers' designs for routinization is only the first step in understanding how the routines function in practice. We must also examine how employers try to persuade or coerce the relevant actors (in this case, service-recipients as well as workers) to cooperate and how the behavior of these actors, and other contingencies, alter the routines in action. (p. 8)

While managers often require workers to follow the script and fulfill the needs of their customers during service interactions (Bélanger & Edwards, 2013; Leidner, 1993; Sallaz, 2002)—and beyond (Sherman, 2011), customers may lack incentives or intentionally do not follow the scripted interactions for their interests (e.g., Bolton & Houlihan, 2005). As such, routinization of service interactions requires the management of workers and customers.

Korczynski (2004, 2007, 2009b) put forward the concept of "customer-oriented bureaucracy" to describe how labor control is exercised through the dual logics of bureaucratic measurement and customer-related norms in service workplaces. Service workers are required to complete a task efficiently, while maintaining a high quality of customer service. Managers often rationalize the authority of customers (Bélanger & Edwards, 2013; Bishop & Hoel, 2008; Fuller & Smith, 1991; Leidner, 1993), as evident by an infamous marketing slogan "The

customer is always right." This can also be considered a form of normative control in which workers are incited to engage in emotional labor to meet the needs of their customers. Yet, recent studies also draw attention to contextual factors that influence worker-customer relationship (Korczynski, 2009a, 2013; Lopez, 2010). For example, depending on the types of emotional relationship (instrumental or caring), power-relations (high customer power or egalitarian relationship), and repetition of service interactions (short-term or long-term), workers can find their interactions with consumers fulfilling, instrumental, or alienating (Korczynski, 2009a, p. 957).

Another line of research focuses on the role of technologies in shaping the power dynamics in service workplaces, particularly in call centers (e.g., Bain & Taylor, 2000; Callaghan & Thompson, 2001; Ellway, 2013; Woodcock, 2017). Woodcock (2017) argues that the integration of telephones and computer-enabled call centers accelerates the labor process and expands the monitoring of individual workers' performance and the scope of data collection. Each phone call is recorded and stored for protecting the interests of call centers and assuring the quality of services. As he describes, "Every call encounter is preserved for eternity, every mistake might be punishable in the future" (p. 66). Additionally, call centers rely on normative control through value-based team training exercises (Fleming & Sturdy, 2011). Surveillance, however, elicits worker resistance. Many forms of worker resistance are relatively hidden and expressed in a covert manner, in part because of the powerless position of call center workers (Woodcock, 2017). Examples of worker resistance include humor (Taylor & Bain, 2003) as well as "cheating, work avoidance, absence and resignation" (Woodcock, 2017, p. 98; see also Mulholland, 2004). From

the perspective of the management, these practices may be considered "misbehaviors." However, these moments of resistance temporally enable workers to disrupt the labor process and gain a sense of autonomy (Woodcock, 2017; see also Ticona, 2015). In this project, I follow this conceptualization of resistance to focus on the moments where workers attempt to claim their agency in their daily work.

Overall, this line of research informs my analytical focus on the interplay between platform owners, workers, and customers. Situating in this three-way contest for control, metrics can be considered managerial efforts to coordinate both workers and customers. On the one hand, metrics provide employers tools to not only monitor workers' performance (Fuller & Smith, 1991), but also manage consumers' complaints. When consumers were dissatisfied with service interactions, managers could simply direct them to make a complaint by giving workers a low rating. On the other hand, the uses of metrics may attribute the responsibility of customer management from employers to workers. As the managerial role of platform owners is much more invisible than managers in traditional service work, this raises questions about the tensions between labor control and resistance in the gig economy.

Algorithmic Labor Control and Resistance in the Gig Economy

The three cases of this project, ride-hailing platforms, TaskRabbit, and delivery platforms, are part of the larger gig economy. A defining characteristic of these labor platforms is the role of algorithms—which are defined as "encoded procedures for transforming input data into a desired output, based on specific calculations" (Gillespie, 2014, p. 167)—in distributing work tasks and managing distributed workforces (Vallas & Schor, 2020). The uses of performance metrics are considered

an essential aspect of algorithmic labor management (Gandini, 2019; Lee et al., 2015; Rosenblat, 2018a; Rosenblat & Stark, 2016; Ticona et al., 2018; Stark & Levy, 2018; Wood et al., 2019), or what Moore and Joyce (2020) call the platform management model (PMM). PMM identifies several common aspects of labor control, including algorithmic allocation and tracking of work, uses of customer ratings, classification of independent contractors, extraction of commission on every service transaction (p. 930). Specifically, by classifying gig workers as independent contractors, labor platforms disavow employment obligations (van Doorn, 2017) and shift work-related risks to workers (Bieber & Moggia, 2020; McDonald et al., 2020; Gregory, 2021). In a study of food couriers in Scotland, Gregory (2021) found that workers bear physical, financial, and epistemic risks at work. Meanwhile, platforms rely on its algorithms and user agreements to govern and prescribe workers' expectations and behaviors (McDonald et al., 2020; Shestakofsky & Kelkar, 2020).

Existing studies have examined how the design and construction of algorithms facilitate labor management of on-demand delivery workers (Cant, 2020; Gregory, 2021; Griesbach et al., 2019; Shapiro, 2018; Veen et al., 2020; Wahl, 2020), ride-hailing drivers (Lee et al., 2015; Mosseri, 2020; Rosenblat 2018a; Rosenblat & Stark, 2016), careworkers (Ticona & Mateescu, 2018), and workers on microwork platforms (Gray & Suri, 2019; Lehdonvirta, 2018; Irani, 2015), among others. Accordingly, algorithmic control becomes what Kellogg et al. (2020) call "the new contested terrain between employers and workers" (p. 372). There are three key elements of algorithmic labor control, namely, algorithmic direction, evaluation, and discipline (Kellogg et al., 2020; see also Edwards, 1984). Algorithmic direction relies on the restriction of work-

related information and personalized recommendations to direct work tasks. An underlying premise is that the control over work-related information is central to a platform's labor management as well as workers' decision-making capabilities. Workers' autonomy is dependent upon what and how much work-related information is available to them through the platform interface (Curchod et al., 2020; Rosenblat & Stark, 2016; Shapiro, 2018). Hence, platforms can govern workers through the selective revelation and withholding of information from workers.

Algorithmic evaluation entails constant monitoring and assessment of workers' performance. The uses of performance metrics are essential to this mechanism of labor control, which focuses on algorithmic surveillance, or what Christin (2018) calls "*quantification as discipline*" (p. 1386; italics original). Metrics can facilitate platform surveillance, as workers are being continuously monitored and evaluated by algorithms and their customers. Following our previous discussion of the reactivity of metrics, when workers are aware of being monitored—whether by the platform or customers—they feel anxious, self-monitor themselves, and modify their behavior in response to what is perceived to be measured. In fact, algorithmic evaluation is closely associated with rewards and sanctions (i.e., algorithmic discipline). Workers, therefore, are incited to score well to earn rewards and reduce the risk of being punished (e.g., warning, deactivation, and other effects on employment opportunities).

Labor control and resistance go hand in hand. Gig workers always find ways to navigate and resist algorithmic labor management, at individual and collective levels. In a study of remote gig workers in Africa, for example, Anwar and Graham (2020) found that workers may exercise agency through practices of resilience (e.g., buying

reviews), reworking (e.g., using multiple accounts and monitors), and resistance (e.g., filtering clients and jobs) on and off the platform. While platforms can use recommendations and offer incentives to encourage when and where workers should work, ride-hailing and delivery workers may simply ignore such algorithmic direction and set their own schedule to strategize their efforts (Rosenblat, 2018a; Rosenblat & Stark, 2016; Veen et al., 2020). Others may proactively use bots and multiple mobile phones to game the algorithms (Chen, 2018). During service interactions, workers may also actively manage their customers by performing emotional labor (Raval & Dourish, 2016; Rosenblat, 2018a; Rosenblat & Stark, 2016) and initiating direct and personal communication with customers (Curchod et al., 2020) to maintain their autonomy. In short, as workers become familiar with the platform's policies and socio-technical features, they may feel empowered and develop their own calculative practices (Shapiro, 2018).

Alongside the day-to-day contestation of algorithmic labor control, gig workers create informal communication networks for exchanging work-related information with their peers on online forums and social media platforms (e.g., Maffie, 2020a; Gray et al., 2016; Sutherland et al., 2020; Ticona et al., 2018; Wood et al., 2018). At the collective level, there have been strikes and demonstrations against labor platforms (e.g., Cant, 2020; Moore & Joyce, 2020; Lei, 2021; Tassinari & Maccarrone, 2020; Woodcock & Graham, 2020). These modes of individual and collective resistance can potentially create "moments in which algorithms do not govern as intended" (Ferrari & Graham, 2021, p. 13), though gig workers meanwhile are at risks of being punished by platform owners.

As this section shows, there are already a number of studies on algorithmic labor control and resistance in the gig economy. In *Uberland*, Rosenblat (2018a), for example, presented an extensive ethnographic study of Uber and Lyft drivers in North America and identified the key mechanisms of algorithmic labor management (e.g., information asymmetries, surge pricing, and ratings). There are also comparative studies of labor platforms. Ravenelle (2019) highlights the precarious nature of gig work in the contexts of Airbnb, Uber, TaskRabbit, and Kitchensurfing. Studying workers on thirteen platforms (e.g., Airbnb, TaskRabbit, Uber, Lyft, Postmates, and Stocksy), Schor's (2020) recent book problematizes the term "sharing economy" and discusses the extent to which these platforms fail to realize their promises. Indeed, all of these studies provide valuable insights into the theoretical and analytical framing of this project. Nonetheless, this project has a narrower focus on algorithmic metrics and evaluation in the context of service work. This contributes to understanding how metrics can mediate and reshape the service triangle on and off labor platforms. Moreover, with a few exceptions (Gray & Suri, 2019; Ravenelle, 2019; Schor, 2020; Wood et al., 2019), many of the existing studies focused on a discrete labor platform. This project, therefore, can offer a comparative perspective of algorithmic labor control and resistance by examining workers' convergent and divergent reactive practices towards metrics.

Research Questions

The over-arching research question of this dissertation is to explore how metrics become sources of algorithmic labor control and resistance across the three cases. Metrics as a form of evaluation that become intertwined with algorithmic

direction and discipline in various ways. One important dimension of algorithmic evaluation has to do with the visibility of metrics. Specifically, how do labor platforms use metrics to govern their distributed workforces? What are the kinds of metrics involved in platform-based labor processes and with what purpose? How do labor platforms legitimize their uses of metrics as part of the labor processes? An examination of these questions contributes to understanding the role of metrics in enabling algorithmic labor control in the gig economy.

Drawing on the literature around the power and reactivity of metrics, the second set of questions has to do with how gig workers encounter, interpret, and respond to metrics. Indeed, there are different kinds of metrics (e.g., customer ratings, acceptance rates, and cancellations rates), with various degrees of visibility to workers and customers across labor platforms. Customer ratings, for example, represent customers' assessment of service quality. Acceptance rates and cancellation rates, by contrast, are system-generated behavioral measures. Do they exercise the same kind of control over workers? How do gig workers interpret and manage different metrics? Meanwhile, the three case studies represent different types of digitally-enabled service work which have their distinct labor processes. In these cases, metrics can directly or indirectly affect workers' employability. Customer ratings on ride-hailing and delivery drivers can determine their continuity of employment on the platforms, whereas such ratings affect the visibility of workers' online profiles on TaskRabbit. Workers' metrics also become intertwined with their qualification for platform-based rewards programs. To what extent are workers' interpretation and management of metrics similar and different across the three cases? What is the potential for labor resistance

across the labor platforms? Overall, this project considers metrics as an emerging form of evaluation in the contexts of digitally-enabled service work. By comparing workers' interpretation and management of metrics across the three cases, this project can deepen our understanding of how metrics shape the trilateral relationship between platform owners, service workers, and customers in the gig economy.

CHAPTER 3

METHODOLOGY

This dissertation is a qualitative multi-case study of digitally-enabled service work in the gig economy in the United States. In this chapter, I start by defining the case study approach and explaining the logic of casing. I argue that this comparative approach is particularly useful for studying workers' practices around algorithmic metrics across labor platforms. Then, I describe the process of data collection, including issues about access, recruitment, and the interview process. I also document the challenges that unexpectedly shaped the process of data collection, and therefore, the interview sample. After that, I discuss my general approach to data analysis. Finally, I reflect on the challenges and successes throughout the research process.

A Multi-Case Study Approach

Before discussing the case study approach, it seems important to point out methodological issues in studying social construction of and practices around algorithms. Drawing on the work of Burrell (2016), Christin (2020b) identifies four dimensions of algorithmic opacity: intentional secrecy, technical illiteracy, unintelligibility of machine learning, and the massive amount of data that algorithms process. Both users and researchers, therefore, may have difficulties in knowing algorithms. Accordingly, this raises a key methodological issue: how might we study something that is seemingly "black-boxed" and inscrutable (Christin, 2020b; Kitchin, 2017; Seaver, 2017; Ziewitz, 2016)? To answer this question requires an understanding of (a) what an algorithm is (Ziewitz, 2016) and (b) how we "know" that we are studying the impacts of the algorithm. Therefore, scholars have advocated for

ethnographic approaches (Christin, 2020b; Seaver, 2017) and mixed-methods approaches (Kitchin, 2017) [9] to unpack socio-technical assemblages of algorithms. Analytically, scholars have turned to users' "algorithmic imaginary" (Bucher, 2017) or "folk theories" (Toff & Nielsen, 2018; Ytre-Arne & Moe, 2020). These analytical approaches focus on how users construct social representations of algorithms (Christin, 2020b). From the perspectives of platform owners or designers, such representations may not be the "correct" understandings of algorithms. However, these representations can drive users' practices around algorithms.

This project, similarly, analyzed how gig workers encountered, interpreted, and managed algorithmic metrics in their daily work. Indeed, algorithmic metrics are objects that are visible to the interviewees. That said, the kinds of metrics I studied in this project are generated and processed by platform-based algorithms. In the three cases, workers can view their metrics in their platform profiles. Platform owners sometimes send emails and in-app messages about the tips for scoring a 5-star rating. In this sense, the objects of analysis seemed clear to the participants, though they might still have difficulties in understanding the inner workings of algorithms.

To study workers' reactive practices towards metrics, I adopted a multi-case study approach. The case study approach is useful for generating an "in-depth understanding of a single or small number of "cases," set in their real-world contexts"

[9] Here, mixed-methods approaches do not mean the combination of using quantitative and qualitative methods. Instead, Kitchin (2017) outlines six methodological approaches for studying different aspects of algorithms (e.g., examining source code, interviewing designers, conducting ethnographies and discursive analysis, etc.). Kitchin, meanwhile, proposes that these approaches have their strengths and weaknesses, so it is useful to employ "two or more of the approaches in combination" (p. 22).

(Yin, 2012, p. 4; see also Eisenhardt, 1989; Schwandt & Gates, 2018). This approach draws on multiple sources of data—which can be qualitative or quantitative—to better understand "how" and "why" a social phenomenon works in a particular context (Yin, 2012). Following this approach, I relied on multiple sources of materials to explore the tensions between platforms' corporate discourses and workers' practices within wider socio-technical assemblages (Kitchin, 2017).

A case is the unit of observation, which can be "a person, an organization, an event, a decision, an action, a location like a neighborhood, or a nation-state" (Schwandt & Gates, 2018, p. 600). Cases can be found or made: the former refers to an empirical and bounded unit, whereas the latter considers cases as theoretical constructs (Ragin, 1992). Regardless of the units or assumptions about cases, researchers must consider the logic of *casing*, which refers to an iterative process of defining and bounding their cases (Ragin, 1992, 2009). To compare the findings of each case, cases must be "similar enough and separate enough to permit treating them as comparable instances of the same general phenomenon" (Ragin, 1992, p. 1).

When I began working on this project, I decided to examine the cases of ride-hailing platforms, TaskRabbit (as an example of temp work platforms), and platform cooperatives (co-ops). The first two cases were chosen to maximize the variation and explore the impacts of labor processes and socio-technical features on workers' interpretative practices around metrics. The case of platform co-ops was intended to serve as a negative case because many of them intentionally do not use customer ratings. Yet, as I will discuss in my research process, I was unable to recruit any interviewees who had worked for platform co-ops. Therefore, I adjusted my casing

logic to select the three cases based on their types of digitally-enabled work (ride-hailing, temp work, and delivery platforms) and the visibility of platform-based metrics. On ride-hailing platforms, workers' ratings are visible to themselves and the riders who request their services. Other forms of metrics remain only visible to workers. On TaskRabbit, workers' metrics are visible to all prospective clients on the platform. On delivery platforms, the visibility of workers' metrics is similar to that of ride-hailing drivers, but DoorDash and Instacart do not allow drivers to rate their customers. In short, the three cases are "positive cases" (i.e., using metrics for labor management), but they are varied in terms of their algorithmic matching processes, the visibility of metrics, and metrics-related rewards and sanctions (see Table 1 in Chapter 1). The three cases, therefore, are similar because of the nature of service work. But they are also different enough to capture how different socio-technical design of labor platforms can influence workers' reception of metrics. In this sense, the cases were chosen "on the basis of expectations about their information content," with an aim of maximizing the variation (Flyvbjerg, 2006, p. 230). Although this project did not use an ethnographic approach, it echoed the strategy of "algorithmic comparison" that "uses a similarities-and-difference approach to identify the distinct features of the technical instruments and their related uses" (Christin, 2020b, p. 912).

In each of the three cases, I drew on multiple sources of empirical materials to make sense of workers' uses of metrics. First, I primarily relied on the interview data to understand how gig workers encountered, interpreted, and used metrics. As it was difficult for me to directly observe workers' actual practices in a rather naturalistic setting, this method allows for exploring participants' cultural frames of their practices

(Lamont & Swidler, 2014). Second, I analyzed discursive texts related to the platforms I studied. This involved the company websites and blog posts and popular press articles. The primary objective of the discursive analysis was to understand how platform owners narrated and legitimized algorithmic labor management. This strategy also helps to compare gig workers' practices with platforms' discourses. Third, I drew upon the discussions from worker-to-worker online communities to supplement the analysis and provide background to the interview data. As I will discuss below, I encountered these communities during the process of data collection. These communities are important social spaces for workers to share work-related knowledge, complain about workplace conditions, and discuss platform policies (e.g., Maffie, 2020a; Gray et al., 2016; Ticona et al., 2018). Taken together, interrelating multiple sources of data can help to explore and complement overlapping yet distinct facets of algorithmic labor control and resistance (Greene et al., 1989; Small, 2011).

I largely followed an interpretive orientation to the case study approach (Schwandt & Gates, 2018). That said, the research process, including the moments of interviewing and analyzing data, was a form of social construction (Denzin & Lincoln, 2011; Fortner & Christians, 1989; Kvale, 2002). I aimed to explore participants' understandings of metrics from their perspective, rather than to generalize the findings from the case studies.[10] I mainly recruited participants via social media platforms (e.g., Twitter), worker-to-worker blogs, and Reddit as well as using a snowball sampling method via participants' networks. This sampling strategy potentially limited

[10] Scholars who adopted a realist approach would focus on generalizing causal mechanisms and effects from an individual case or multiple cases (Schwandt & Gates, 2018; see also Flyvbjerg (2006)'s discussion of the misunderstandings about case study research).

the sample to those who were more vocal about the labor platforms. Nonetheless, these groups of participants were arguably sensitive to the workings of the platforms. I received IRB approval for this project (see Appendix 1-1 to Appendix 1-8 for the recruitment materials and interview guides) before recruitment and data collection. I also assigned pseudonyms to anonymize the identities of the interviewees to comply with the IRB's requirement.

Overall, I conducted in-depth interviews with 50 gig workers in the United States, including 21 ride-hailing drivers, 11 TaskRabbit Taskers, and 18 delivery workers. While I collected the majority of the interview data for the case of ride-hailing platforms between 2017 and 2018, I conducted interviews with workers for the other two cases in 2020.

Data Collection as a Research Process: Access, Recruitment, and Interviews

In qualitative research, the researcher is the key instrument to gather and interpret "relevant" information (Creswell, 2014); therefore, it is vital to explicate the research process and methodological choices in detail (Humphreys, 2020; Roy et al., 2015). Indeed, this project did not go exactly as I had planned. For example, I had originally planned to include a "negative" case—service worker-owned platform co-ops which explicitly do not use any forms of ratings—to "offer theoretically decisive contrasts with positive cases" (Ragin, 2009, p. 525). I attended the 2019 annual consortium of platform cooperativism in New York City and found it a great resource for understanding the principles and practices of platform cooperatives. Unfortunately, however, I was not able to recruit any participants who worked for a platform co-op. Therefore, I eventually decided to study delivery platforms as my third case.

Additionally, I originally planned to recruit at least 20 participants for each case (in total 60 participants) and potentially conducted my fieldwork in New York City, one of the major cities for platform businesses. However, the COVID-19 pandemic has exposed gig workers to heightened economic and health risks. It also affected my research and travel plan. I raised these unplanned issues here not only to highlight the constraints, but also to consider how these constraints might shape the methodological choices I had made.

In what follows, I will discuss issues about the study populations (when the data were available), access, recruitment, and interviews for each case. It is noteworthy that access and recruitment occurred almost at the same time because there were no physical workplaces or field sites. The project can be roughly divided into two phases. I began by studying ride-hailing platforms in the early phase of my research in 2017 and 2018. I continued recruiting interviewees who were Uber and Lyft drivers in January and November 2020. Additionally, I collected the interview data for the cases of TaskRabbit and delivery platforms between September and December 2020.

Ride-hailing Platforms: Uber and Lyft

Uber and Lyft are the two major ride-hailing platform companies in the United States. There were other local ride-hailing companies such as Juno in New York City between 2016 and 2019, but the interviewees for this project worked for Uber and/ or Lyft. By February 2021, Uber's ride-hailing service is available in 265 cities, whereas Lyft's service is available in over 600 cities in the United States. Yet, Uber has a stronger global presence and a larger market share (71%) than Lyft (29%) in the

United States (Yeo, 2021). There are no official statistics about the demographics of workers who work for Uber and Lyft. It was estimated that Uber had about 750,000 active drivers, whereas Lyft had about 700,000 active drivers in the United States in 2017 (Rosenblat, 2018a). Uber (2020) claimed that one million Americans used Uber to generate income in 2019, but this number included those who offered ride-hailing services and delivery services (i.e., Uber Eats).

The two platforms also did not clearly define the term "active drivers." Lyft's (2019) IPO prospectus estimated that 91% of the drivers worked fewer than 20 hours per week. Based on the administrative data from Uber in over 20 U.S. cities from 2012 to 2014, Hall and Krueger (2018) found that 83% of drivers drove fewer than 35 hours a week. The majority of the working population was male (86.2% of the respondents) and in the age range from 30 to 64 (78.2%). They also found that Uber drivers' average net hourly income (before vehicle expenses) was $19.35, as compared to taxi drivers ($12.56). However, Berg and Johnston (2019) have criticized that the calculation did not account for self-employment taxes, licensing costs, location-specific mileage costs, vehicle ownership costs, and the changes in Uber's fare (pp. 47-48). There were other unofficial estimates. Since 2016, "The Rideshare Guy," one of the most popular blogs about the ride-hailing industry, collected and published survey data from Uber and Lyft drivers who subscribed its email list on an annual basis. Its 2019 survey suggested that Uber and Lyft drivers made $13.47 and $11.55 per hour after expenses, respectively (Campbell, 2021). Drivers who drove for both Uber and Lyft reported making $12.66 per hour after expenses. This survey did not use a random sampling method. The lack of official data about gig workers also

represents a common challenge for studying the gig economy (Klawitter, 2020).

Access and recruitment. Although I did not aim to recruit a "representative" interview sample, gaining access remained a key challenge. This is in part because recruiting participants on the platforms could potentially violate the terms of service (see also Humphreys (2020)'s discussion of the Bumble study). When I first started my research on Uber drivers in December 2017 and the early 2018, I tended to recruit participants during my Uber rides. At the time, I was interested in understanding Uber drivers' performance of emotional labor (see Appendix 1-1 for the interview guide). I initially planned to conduct interviews with Uber drivers during a ride, if they agreed to do so. However, there were concerns from IRB that conducting interviews during a ride might distract drivers from driving. In order to comply with IRB's requirements, I eventually recruited drivers during rides, but then set up another time to interview them later when they were not working.

This strategy, however, did not work quite well. When I tried to recruit drivers for the interviews, they often wondered why the interview could not be done during the ride. Although a few drivers agreed to be interviewed and gave me their phone number, only one interview was scheduled. There are two possible reasons. First, as I did not provide any compensation for their time, interviewing took time and placed an extra burden for them. Second, though I mentioned that the research would not affect the ride during the recruitment (see Appendix 1-2 for in-person recruitment message), it was possible that drivers were under pressure to accept my interview request because I was the customer during the ride.

Then, I changed my research strategy to recruit drivers in driver-to-driver

online communities, in part because I learned that drivers share work-related information and strategies online from a news article by Rosenblat (2018b). YouTube was one of the sites where I tried to recruit participants. In early 2018, I used a keyword search (e.g., "Uber" and "Ridesharing") on YouTube's search engine. Then I relied on YouTube's recommender system to do a snowball sampling to identify more YouTubers who shared content about Uber. I looked at the profiles and videos of these channels (n = 61) and excluded 26 of them because they had not published videos for more than a month and did not include a description of the channel. Many of the sampled channels posted email address in their description; therefore, I mainly used e-mail to solicit participants (see Appendix 1-3 for the sample recruitment email). In addition to YouTube, I contacted Uber drivers who blogged about their experiences to see if they were interested in being interviewed. Nine of the interviewees were recruited through publicly found information on their YouTube channels or blogs. Some of them shared my research to their subscribers and viewers. I also asked whether the interviewees could introduce other participants to me, using a snowball sampling method (Humphreys, 2020). A few of the interviewees gave me email addresses of their friends or subscribers who might be willing to chat with me. I followed up with email to arrange interviews with two participants.

In January 2020, I planned to recruit more participants for my dissertation. But at this point, I decided to offer a $20 gift card to compensate for the time of participants. I posted a recruitment message on Twitter (see Appendix 1-4 for the screenshot of the recruitment material) and reached out to the driver/bloggers to see if they could spread the information about my study. Harry Campbell, the founder of

"The Rideshare Guy," shared the information about my research with his followers on Twitter. I scheduled five interviewees after posting the message. From November 18 to December 1, 2020, I posted five Reddit advertisements to recruit participants who worked for Uber, Lyft, TaskRabbit, Instacart, or Uber Eats. I will discuss the Reddit advertisements later in the next sub-section because I only recruited two Uber or Lyft drivers via Reddit (see Table 2 for the number of members of the subreddits where I posted an advertisement). I stopped the advertisement on December 1 because no Reddit users contacted me about my research in this week.

Table 2. The Number of Members of the Subreddits

Subreddit	Members of the Subreddit
r/uberdrivers	135,000
r/lyftdrivers	19,200
r/TaskRabbit	2,200
r/UberEATS	30,000
r/InstacartShoppers	46,400

Interview Data Collection. I interviewed a total of 21 Uber and/or Lyft drivers. 14 of them worked for Uber and Lyft. All of the interviews were semi-structured (Stewart & Cash, 2003). The main topics included participants' background, uses of the platform in general and the rating system in particular; interactions with their customers; acquisition of work-related information; and reflections on their work. As I did not have access to the worker version of the platforms I studied (except for the screenshots the platforms posted on their websites), I also asked participants to walk me through the app they used during the interviews. As previously discussed, I used different versions of the interview questions at different stages of the research (see Appendix 1-1, Appendix 1-5, and Appendix 1-6**)**.

For example, when I conducted interviews with those who created content on YouTube or their blogs (9 of 21),[11] I included questions about their performance of expertise and self-branding activities. Nonetheless, a major part of the interview protocol was about their Uber work experience. The interview guide I used for the dissertation project (see Appendix 1-5) was adapted from the previous project about Uber drivers' performance of emotional labor. Therefore, most of the questions were consistently asked throughout the research period.

Before the interview, I checked with the participants whether they would prefer using phone, Skype, or other software (e.g., Zoom) for conducting the interview. I sought participants' oral consent to participate in my research and audio-record our conversation at the beginning of the interviews. At the end of each interview, I asked them if they could introduce anyone who might be interested in chatting with me. Finally, I sent an email to thank the participant and included the information sheet about my research for their reference.

The majority of the interviews (14 of 21) were conducted between 2017 and 2018, whereas the rest of them were conducted between January and November 2020. The interviews lasted from 18 minutes to 2 hours. As I mentioned before, I began offering compensation for participants in 2020; therefore, seven interviewees received a $20 gift card after the interviews. Table 3-1 in the Appendix offers a brief profile of

[11] I interviewed 11 driver/bloggers who drove for Uber and created content about the ride-hailing industry on YouTube or their blogs and published a paper about this topic. Unfortunately, however, I had to exclude two of the interviews from this project because the interview recordings and transcripts were lost. I stored the files in my old laptop, but I could not find them as I began to re-look at the files for conducting the analysis in November 2020. I could only find the email and Facebook conversations I had with the two participants.

the ride-hailing drivers I interviewed. The sample included 17 men and 4 women. Geographically, eight of them worked in California (Los Angeles, San Diego, Monterey, and San Mateo). The rest of them worked in Chicago (n = 4), Syracuse (n = 2), Ithaca (n =1), Louisville (n = 1), Charlotte (n = 1), Miami (n = 1), Baltimore (n =1), Georgia (n = 1), and Washington, DC (n = 1). The length of working on the platform(s) varied from 3 months to about six years. Their average weekly working hours also varied. For example, Lincoln and Timothy did not have a stable working schedule for Uber and Lyft, as they had a full-time job. Daniel, Noah, and Clayton told me about the typical working hours per day. For others, the average weekly working hours ranged from 10 hours to 72 hours. I did not include age into the table because I did not ask about the participants' age during the majority of the interviews (except for delivery workers).

Temp Work Platform: TaskRabbit

TaskRabbit is a temp work platform where individual clients and organizations can hire Taskers for over 45 categories of tasks such as moving, cleaning, home repairs, furniture assembly, and personal assistants. Since IKEA acquired TaskRabbit in 2017, the platform's furniture assembly service is available in IKEA stores in the United States and other key markets in Canada and Europe. The platform currently operates in 53 cities in the United States. The most popular categories of tasks, according to TaskRabbit's (2020) "Fact Sheet," are furniture assembly (1.3 million over the last decade), moving (825,000), home cleaning (579,000), and home repairs (438,000). TaskRabbit states that over 4 million tasks were completed over the last decade. Like the other two cases, TaskRabbit does not frequently release the official

data about workers' demographics. In 2013, the company reported that 70% of the Taskers had a Bachelor's degree, 20% had master's degrees, and 5% had a PhD (Newton, 2013). By October 2018, TaskRabbit (2018) claimed that "there are **more than 140,000 Taskers** on the network. **60%** are Millennials" (emphasis original). In January 2020, TaskRabbit (2020) estimated that Taskers' average hourly earnings were $36 in the United States. Crawling data from all worker profiles (N = 3,707) on TaskRabbit in December 2015, Hannák et al. (2017) identified gender and race based on profile images. 88% of the profiles were identified. Women accounted for 42% and men accounted for 58%. In terms of race, 73% were White, 15% were Black, and 12% were Asian. The demographics data only served as a reference point as they were crawled in 2015.

Access and recruitment. The recruitment process for Taskers was similar to that for ride-hailing drivers. I mainly recruited participants from worker-led online communities and snowball sampling. While I primarily looked at YouTube for recruiting ride-hailing drivers, there were very few active YouTube channels for Taskers. I also did not recruit any participants during the tasks, primarily because TaskRabbit was not available in Ithaca, and even it was available, face-to-face communication would be significantly limited during the COVID-19 pandemic. My recruitment strategies were also slightly different from existing interview-based studies about Taskers. Schor (2020), for example, recruited Taskers by posting the interview as a task on the platform between 2013 and 2015. Nonetheless, TaskRabbit has changed its labor process since July 2014, meaning that it no longer allows me to post a task publicly for Taskers to bid on the task. Instead, I could only request a

specific Tasker and hire them for the task. “Research” is also not a common task category (i.e., usually under the category of “Personal Assistant”).

Therefore, I recruited my participants *outside* of TaskRabbit. First, I registered a “Recruiter” account on LinkedIn in September 2020. This allowed me to send “InMail” messages to people whom I did not have any connections on LinkedIn. This function, however, only enabled me to contact those who chose to receive an “InMail” message. Then, I used LinkedIn’s search function to look for people who claimed themselves as a current or former Tasker on their profiles. In the “InMail” message, I simply stated that “I’m conducting a study on how Taskers use mobile technologies and interact with clients” and asked whether they would be interested in participating in the study (i.e., similar to the sample recruitment material in Appendix 1-3). I was able to conduct an interview with Kuzma, and he later posted my research in some of the closed Taskers-related Facebook groups. This helped me to expand my interview sample, using a snowball sampling method.

Second, as I mentioned previously, I posted a series of Reddit advertisements on different subreddits from November 18 to December 1, 2020. Appendix 1-7 showed the sample of the advertisements for participant recruitment. I used “Looking to interview Taskers for a paid research project (Remote interview)” to incentivize participation. Depending on the target groups and subreddits, I slightly changed the title and content (e.g., from “Taskers” to “Uber Eats drivers). The users who were interested in participating in the project contacted me either by replying to the posts or sending me a private message on Reddit. I then followed up with them to describe the interview process and schedule the interview time.

Reddit allowed me to define the target audience and specify when to show the advertisements. I chose to focus on the United States and set the "Communities" to a specific subreddit (see Appendix 1-8 for an example). The advertisements were pinned posts and available from 9:00 am to 11:00 pm in the selected period. This strategy was more successful in recruiting delivery drivers rather than Taskers and ride-hailing drivers. I was only able to recruit two Taskers from Reddit. However, one of them did not work in the United States, so I excluded him from the project. I failed to recruit many Taskers via Reddit, possibly because the subreddit had a much lower number of members, compared to other subreddits I posted advertisements.

To summarize, though I tried to recruit participants via LinkedIn and Reddit, the interview sample for this case was primarily based on snowball sampling (9 of 11). The participants knew about my research either from their friends or Kuzma's and others' social media posts.

Interview Data Collection. In total, I conducted semi-structured interviews with 11 Taskers (see Table 3-2 in the Appendix for the interviewees' profile). The interviews were conducted over the phone or Zoom. I used a similar semi-structured interview guide for each interview (see my previous discussion on the interview process for the first case and see Appendix 1-6). Because Taskers could set their schedule and hourly rate, I asked them how they created online profile, managed their schedule (e.g., back-to-back appointments), and set their rate. As the interviews were conducted during the COVID-19 pandemic, I also added a question "How does the pandemic affect your work in general?" during the interview. After each interview, I provided them a $20 gift card to compensate for their time.

I conducted the interviews between September and December 2020. The interviews lasted from 31 minutes to 1 hour and 40 minutes. Although previous study suggested that women accounted for 42% of TaskRabbit's workforce (Hannák et al., 2017), men (n = 8) were overrepresented in my interview sample. The interviewees reported working in Los Angeles (n = 2), Denver (n = 2), Cincinnati (n = 1), Atlanta (n = 1), Dallas (n = 1), Portland (n =1), San Diego (n = 1), Dallas-Fort Worth area (n = 1), and East Coast (n = 1). Their TaskRabbit working experience varied from six months to nine years. Additionally, their average weekly working hours ranged from 6 hours to 45 hours.

Delivery Platform: DoorDash, Uber Eats, and Instacart

DoorDash and Uber Eats are two of the largest on-demand food delivery platforms in the United States. According to DoorDash (2020a), DoorDash and Caviar (i.e., a delivery platform owned by DoorDash) had about 50% of the market share in the United States, whereas Uber Eats had about 26% in October 2020. DoorDash's IPO prospectus explicitly addressed the opportunities and challenges of the pandemic: the pandemic has contributed to its skyrocketing growth rate (i.e., 226% between 2019 and 2020). The actual number and demographics of Dashers and Uber Eats drivers in the United States remain unclear. Uber does not release any data specifically about Uber Eats drivers' demographics and earnings.[12] According to DoorDash, the total number of Dashers has increased from 100,000 in 2017 (in the U.S. and Canada)

[12] Uber discusses the delivery fees for Uber Eats drivers in the UK because the company is required to pay the minimum wage to its drivers. For example, the minimum fee (before Uber service fee) is 3.5 pound in London. But such information is not available in Uber's U.S. website.

(DoorDash, 2017b) to over 1 million (in the U.S., Canada, and Australia) in 2020 (DoorDash, 2020a). In 2019, DoorDash estimated that Dashers' could earn $18.54 on average per hour in October 2019, but this hourly rate included customer service tips (DoorDash, 2019). While the pandemic may increase the demand for delivery services, drivers' pay remains unstable. Drawing on interviews with food delivery workers in New York City, reporter Kimiko de Freytas-Tamura (2020) wrote, "food delivery apps say drivers can earn as much as $22 per hour, including tips, though many drivers said they never earned anywhere close to that much."

Instacart primarily offers grocery delivery and had an explosive growth in its market share during the pandemic (from about 30% in April 2019 to 57% in April 2020) (Damiani, 2020). Before April 2020, Walmart was the leading market player in online grocery delivery market. Instacart's service is currently available in over 5,500 cities in the United States and Canada (Mascarenhas, 2020). There are two types of shoppers in Instacart. Full-service shoppers are independent contractors and account for the majority of Instacart's workforce. Their pay is dependent on the batches and customer service tips. In-store shoppers are Instacart's part-time employees who help to shop and pack the items in a specific store, but not all stores have in-store shoppers. Also, in-store shoppers have a flat hourly rate[13] and do not receive any tips. As I discussed earlier, Instacart has been expanding its workforce to over 500,000 shoppers during the pandemic (Lavietes & McCoy, 2020), but the expansion only applies to

[13] Sociologist Chelsea Wahl (2020) conducted an ethnographic study of Instacart and worked as an in-store shopper at a Wegmans store in a suburb of Philadelphia in 2018. She noted that "in-store shoppers earned $10 per hour with an additional "service bonus" that was typically 5% of each order" (p. 85), but Instacart frequently changes its pay structure.

full-service shoppers. *CNN Business* reported that there were about 12,000 in-store shoppers before the pandemic, but some in-store shoppers reportedly lost their jobs in California, Illinois, Texas, and Chicago (O'Brien, 2020).

Access and recruitment. I mainly recruited delivery drivers via Reddit. As I discussed earlier, I advertised my research on two subreddits where delivery workers shared their work-related experiences (i.e., r/InstacartShoppers and r/UberEATS). I originally had not included DoorDash into the advertisements mainly due to budget concerns. But later I found that almost all of the replies to my advertisements—including the posts on r/uberdrivers and r/lyftdrivers—asked me whether I would be interested in interviewing DoorDash drivers. Indeed, the majority of the interviewees (12 of 18) worked for DoorDash in some capacity (see Table 3-3 in the Appendix for the list of the interviewees). I was able to schedule a number of the interviews and received fewer requests as time went on, so I did not place another advertisement on the subreddit for DoorDash drivers (i.e., r/doordash_drivers). Like the other two cases, I used a snowball sampling method to ask participants for recommendations for other potential interviewees, but I recruited the majority of the interviewees (16 of 18) via Reddit.

There was also one notable instance during the process of recruitment. After posting the advertisement on Reddit, an administrator from a closed worker-to-worker online group invited me to share the information about my research with the members. The online group was designed to provide a space for platform-based delivery drivers (e.g., DoorDash, Postmates, and Uber Eats) to share their stories and chat with others. The administrators were not associated with any of these platforms. After I posted the

recruitment message for members who self-identified themselves as DoorDash drivers, one of the members replied to my message and shared a news article about how Amazon might spy on its workers' online profiles to prevent workers from participating in collective actions. This member warned others that participating in any research interviews might threaten their interest in remaining an independent contractor. As a response, I clarified how I would protect and use the interview data for research purposes only. However, none of the members eventually contacted me later. This instance may reveal gig workers' fear of surveillance.

Interview Data Collection. I conducted 18 interviews with people who worked on DoorDash, Uber Eats, or Instacart (see Table 3-3) between November and December 2020. The interviews were conducted over the phone or Zoom. The interview guide was similar to that in the other two cases, but I modified the wordings to correspond to the platform interfaces (e.g., the rating systems these platforms deploy). The interviews lasted between 11 minutes and 44 minutes.

In the interview sample, only three of the interviewees worked for Instacart. Adele worked as an in-store shopper, whereas the others were full-service shoppers. Therefore, the analysis was primarily based on the interviews with DoorDash and Uber Eats drivers. I conducted interviews with 15 male and 3 female drivers. For this case, I asked the age range of all interviewees. The interviewees' age ranged from 18 to 53. Thirteen of the interviewees were under the age of 30. They worked in Massachusetts (n = 4), Chicago (n = 1), Los Angeles (n = 1), Maryland (n = 1), Pennsylvania (n = 1), New York (n = 1), Central Indiana (n = 1), Mississippi (n = 1), Atlanta (n = 1), Arlington (n = 1), Portland (n = 1), New Mexico (n = 1), Southwest

Virginia (n = 1), Columbus (n = 1), and Washington, DC (n = 1). Their length of working on the platform was between two to three months and two and a half years. Eight of the interviewees began working for DoorDash since March 2020 (i.e., the time when COVID-19 was declared as a national emergency in the United States). Their average weekly working hours varied significantly. Among the interviewees, Sofia spent the least time working for DoorDash. She only worked for DoorDash for 15 minutes to three hours per week. By contrast, Lorenzo worked for Uber Eats and DoorDash for about 35 to 45 hours per week.

Data Analysis

In this project, I followed LaRossa's (2005) framework to analyze the interview transcripts. Drawing insights from the grounded theory (Glaser & Strauss, 1967), LaRossa proposes the principles that guide the three phases of coding, namely, open coding, axial coding, and selective coding. One of the key principles that I found particularly useful was to constantly make "empirical and conceptual comparisons" (LaRossa, 2005, p. 855). Since I started the project about ride-hailing drivers in the early phase of this project, I compared the similarities and differences between drivers' practices and the empirical materials for the other two cases. For example, as Uber uses its rating system to evaluate workers' performance and the quality of service interactions, I considered how the other platforms might evaluate their workforces and frame such technologies of evaluation in similar and different ways. The analysis of the interview data and the discursive texts concerning the labor platforms also provided the comparative basis for thinking through the tensions between algorithmic labor control and resistance.

The analysis began with open coding, which identified emerging indicators (e.g., the key terms used by the participants) and associated concepts (LaRossa, 2005). I read the interview transcripts to look at how the interviewees described their uses of the labor platform(s) and metrics; their interactions with their customers; rewarding and frustrating experiences with the platforms; and their communication with the labor platform(s). Axial coding stresses the relationship between different emerging codes (LaRossa, 2005). For instance, I considered the factors (e.g., information asymmetries, limited communication with platform companies, and the negative ratings they received, etc.) that might contribute to the interviewees' affective feelings of metrics. I also compared how platform-based social-technical features—such as the rewards and sanctions of metrics—might influence drivers' reactions to metrics. The final phase of selective coding helped to craft the main story about how labor control and resistance were connected with the management *by* and *of* customers in workers' daily practices.

During the data analysis, I reflected on the similarities and differences between the interview samples for the three cases. First, all of the interviewees might generally be more vocal about the labor platforms than other workers because they participated in worker-to-worker online communities. Nonetheless, some of the interviewees for the case of ride-hailing platforms were arguably active members of these online communities. They not only sought work-related information but also created content for other workers. Second, the majority of the interviewees were men (40 of 50). There are two possible reasons. First, driving and delivery work are male-dominated industries. Second, female gig workers might be in a more vulnerable position to accept a stranger's interview request in these online communities. The analysis,

therefore, was skewed to the perspectives of male gig workers. Third, the interviewees for the case of delivery platforms might be considered "newcomers" to the gig economy, compared to the interviewees for the other two cases. They started working on the delivery platforms due to the pandemic. This might serve as an analytical point to consider how these newcomers learned to navigate labor platforms' algorithms and metrics.

Methodological Reflection: Challenges and Successes

Before presenting the empirical analysis, I would like to reflect on (a) the methodological challenges I encountered and (b) the ways that my research strategies might influence the validity—which is defined as "a kind of accuracy or correctness" (Humphreys, 2020, p. 81)—of the findings.

One of the major challenges for this project was that gig workers do not have a bounded workplace where I could conduct observational fieldwork and recruit potential participants. The platforms I studied also have limited or no services in the place I live (i.e., Ithaca). While I was able to recruit participants via worker-to-worker online communities, it was difficult to sample gig workers based on their demographics. I also did not include questions about age and income level during most of the interviews; therefore, the analysis did not account for how these factors might influence workers' reception of metrics. To address these issues, I should have included these interview questions and recruited more participants in a major city.

Moreover, as I discussed earlier, the pandemic might result in a higher demand and supply for delivery services. In the case of delivery platforms, many of the interviewees were newcomers. On the one hand, it enabled me to analyze why they

decided to work in the gig economy and how they learned to navigate the work process. On the other hand, these newcomers might not plan to continue working on the platform *after* the pandemic. Therefore, they might not have a strong incentive to dive into platform-based socio-technical features. The interviewees who had a longer working experience tended to provide "richer" data than the others.

There were also constraints of the interview format. All of the interviews were conducted over the phone or video conferencing tools (e.g., Zoom, Skype, and Google Hangouts). This interview format was effective in recruiting participants because the interview could take place anywhere. Nonetheless, the lack of face-to-face communication made it difficult for me to view workers' profiles on the platform and ask follow-up questions. To address this issue, I asked the participants to walk me through the platform as well as their profiles and metrics. This allowed me to gain a better understanding of workers' routine use of the platform.

Overall, there are potential biases from my sampling strategies. The interview sample was skewed to male gig workers and those who were likely to have a stronger opinion about their work experiences. In order to gain a deeper understanding of the labor platforms and workers' perspectives, I also relied on other sources of materials (e.g., news articles and company websites) to complement the interview data.

CHAPTER 4

ALGORITHMIC LABOR MANAGEMENT AND METRICS IN CONTEXT

This chapter situates the role of metrics in algorithmic labor management in the contexts of the platforms I studied. First, I briefly review each platform's development trajectory. The goal is not to offer a comprehensive historical account of the labor platforms, but to contextualize this project. Then, I describe platform-based labor processes; specifically, the steps workers must take to complete an order. Platforms as socio-technical infrastructures enable and mediate workers' experience. As Daniel, an Uber driver, told me, the Uber app "is the Uber experience. I can't do what I do without it." Taskers and platform-based delivery drivers, similarly, need to login their accounts and turn on the app in order to chat with their clients and receive the information about the delivery order, respectively. I also focus on how metrics become intertwined with algorithmic direction and discipline. Then, I offer a synthesis of these platforms' development trajectories and labor processes.

Lastly, I discuss how platform owners legitimize their uses of metrics on a discursive level. Metrics are imbued with an array of meanings, with specific aims. Platform owners framed metrics as boons for the "entire community" rather than a means to manage workers. An inquiry into platforms' discursive work around metrics can help to explore how they articulate the meanings of metrics.

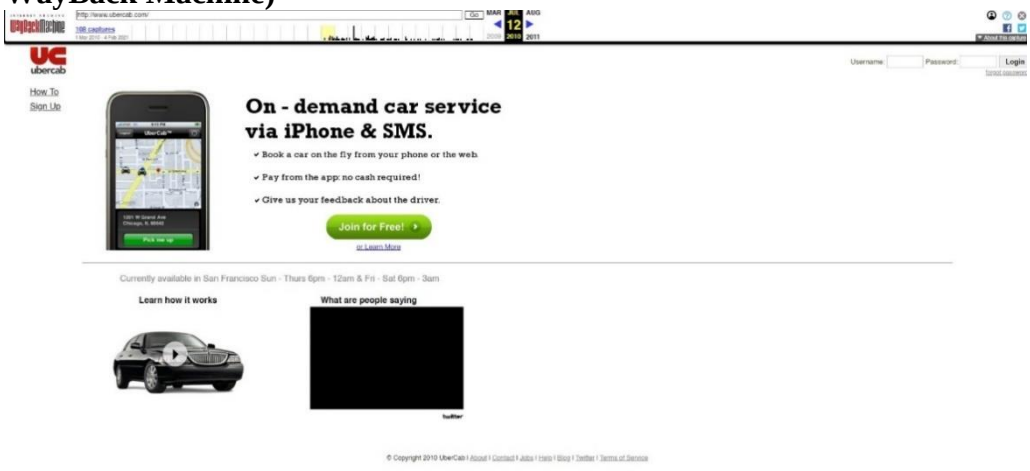
Ride-hailing Platforms: Uber and Lyft

A Brief Overview of Uber's Development Trajectory

In 2008, when Garrett Camp and Travis Kalanick first pitched the idea of "UberCab" (i.e., the original name of Uber), the company envisioned itself as the

"NetJets of car services" and only offered luxury limo-hailing services for professionals who were members of the company (Bhuiyan, 2017). UberCab began its operation in San Francisco in June 2010. Its service costed about 1.5 times more than a typical taxicab (Issac, 2019). UberCab initially had used a website (see Figure 1-1) and soon developed a mobile app for allowing people to request ride-hailing service through iPhone and SMS (Issac, 2019). At that time, UberCab already considered drivers as "partners" to distance itself from employment responsibilities.

Figure 1-1. Screenshot of UberCab's Website on July 12, 2010 (Source: The WayBack Machine)

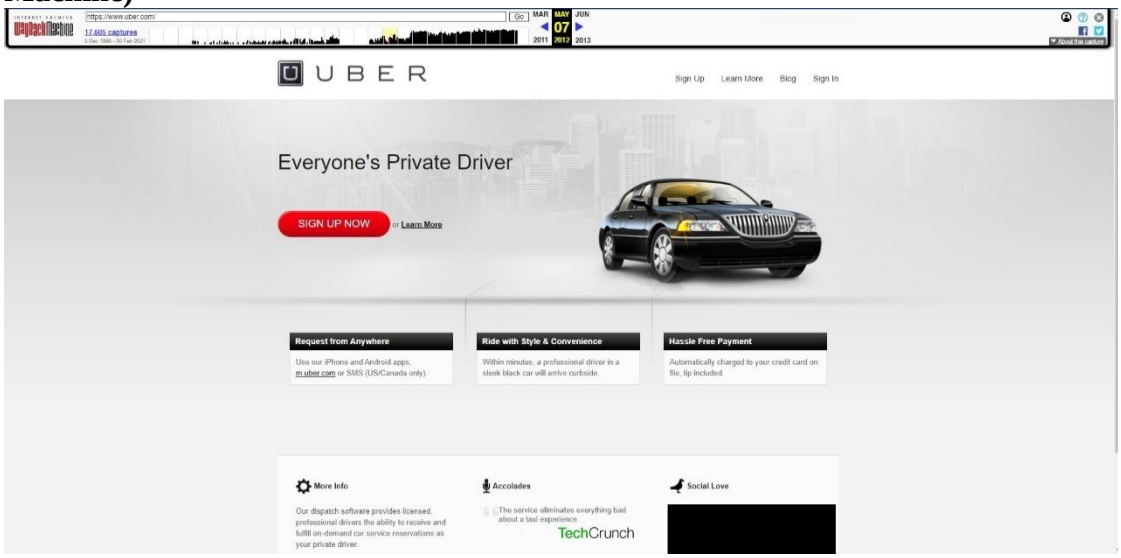


Note. I used a black box to hide the personal information of the people who tweeted about UberCab.

In October 2010, the company renamed itself "Uber," in part because the San Francisco Municipal Transportation Agency denounced the company for offering unlicensed taxi services. According to Issac (2019), Travis Kalanick, a co-founder of the company, stated in the company's office: "We ignore it … We'll drop 'Cab' from our name" (p. 112). Ignoring and challenging the regulations, in fact, has been key strategies for Uber and many labor platforms to expand their businesses (Pollman & Barry, 2017). After the rebranding, Uber began expanding its business to New York in

2011 and other major cities (e.g., Chicago) in 2012.

Figure 1-2. Screenshot of Uber's Website on May 7, 2012 (Source: The WayBack Machine)



Note. I used a black box to hide the personal information of the people who tweeted about Uber.

Figure 1-2 is a screenshot of Uber's website on May 7, 2012 from the WayBack Machine. Uber's motto became "Everyone's Private Driver" and rolled out a cheaper non-limo hailing service called "UberX" which cost 35% cheaper than Uber's luxury service (Tsotsis, 2012). This service significantly expanded Uber's workforce, as it allowed cab drivers and drivers with no business licenses to offer services. In this vein, driving for Uber might become illegal, but the company offered monetary incentives to recruit new drivers in every new market they entered, as evidenced by Craigslist advertisements such as "Let drivers know they get $500 cash when they take their first ride on UberX" (Issac, 2019, p. 19). Recalling his experience driving for Uber in 2014, my interviewee Eddie said,

> I was waiting to pick someone up at union station … My car was in the wrong place. I should have driven it down a different lane, but the person who pulled

> me over once he asked me if I was an Uber driver and I just said yes, and he wrote me a ticket for driving an illegal taxi. I contacted Uber when that happened and they did reimburse me for the cost of the ticket. It definitely scared me when that happened.

Although Uber covered the fine, Eddie's reflection revealed the risk of driving for Uber in its early days.

Alongside Uber's violation of transportation regulations, there have been lawsuits against Uber's employment classification of drivers since 2013 (Hawkins, 2019a). Drivers are supposedly independent contractors, but they are subject to socio-technical control; for example, they cannot see the destination before accepting a ride. Yet, Uber has insisted on its framing and classification of drivers: they are self-employed "driver-partners" and "consumers" of its technology and digital intermediary service (Rosenblat, 2018a). Whenever the government requires Uber to re-classify drivers as employees, or provide certain employment protection, the company threatens to limit or shut down its business.

In 2017, Uber faced a string of scandals such as its false advisements about drivers' earnings, its use of software for identifying police and deceiving law enforcement, and sexual harassment in its workplace (Levin, 2017). In August of 2017, Dara Khosrowshahi became Uber's new CEO to replace Travis Kalanick. Uber, meanwhile, kicked off a campaign called "180 Days of Change" and stated in its email to drivers: "You've told us what you want and now it's time we step up and give you the driving experience you deserve, because simply put, Uber wouldn't exist without you" (Uber, 2017). Of relevance to this project, Uber has redesigned its mobile app to

incorporate a tipping option, ratings protection, and allow drivers to earn "badges" (Hawkins, 2018). Riders are required to select a reason from Uber's list (e.g., "Too many pickups", "Route by Uber" and "Navigation") when they rate a driver below 5 stars. If the selected reason is out of drivers' control, the rating shall be automatically removed from drivers' average ratings. Another relevant change is the launch of Uber Pro program since November 2018. This program strengthens the connection between drivers' ratings and rewards. However, as I explain in Chapter 5, these changes did not seem to address the interviewees' frustrations about ratings.

In early 2021, ride-hailing services remain one of Uber's key businesses, but the company is also expanding its scope of personal mobility market (e.g., e-scooters) and offering delivery (Uber Eats) and logistics (Uber Freight) services.

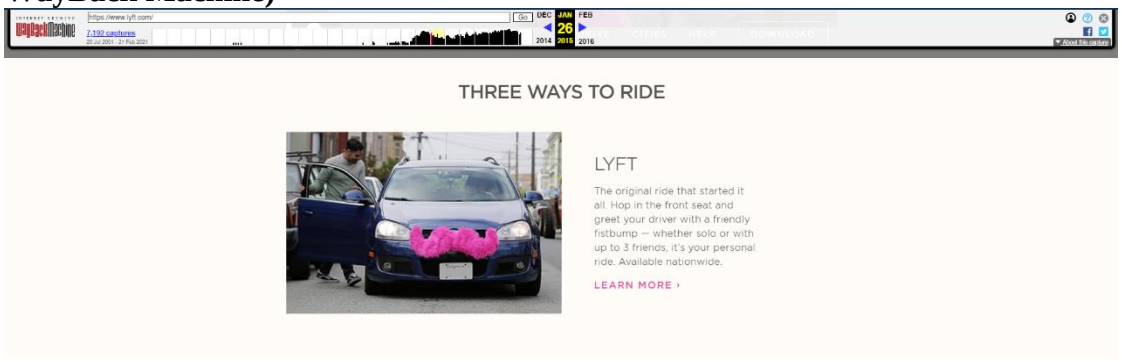
A Brief Overview of Lyft's Development Trajectory

Lyft was initially a "ride-sharing" service offered by a carpooling company called Zimride in 2012 (Cutler, 2012). In 2007, Logan Green and John Zimmer launched Zimride through Facebook at Cornell University (i.e., Zimmer's alma mater). The company aimed to connect drivers with passengers (i.e., college students) who needed long-distance and school-specific carpooling services. As Facebook was the platform where people could request for Zimride, passengers could know drivers' identities. Six months after Zimride was launched, 20% of Cornell students signed up for the service (Schomer, 2011). Later, the company developed its business to different college campuses (e.g., Stanford University) and collaborated with corporations like Walmart (Sullivan, 2009).

In June 2012, Zimride officially started Lyft in San Francisco. What

distinguishes Lyft from Zimrides' carpooling service is that the former focused on short-distance trips. Its marketing campaign was different from Uber's early campaign: while Uber focused on luxury services, Lyft emphasize the vision of building a community (Olanoff, 2012). The vision of community was reflected in Lyft's slogan "Lyft is your friend with a car" and symbolized by a fuzzy pink mustache (see Figure 2) on the front of every Lyft's vehicle. As Zimmer recalled the reason for adopting this iconic symbol in a media interview, "We originally thought of doing this just for women, as a safety kind of service and a very particular clientele. It's partially because of that" (Olanoff, 2012). Compared to Uber, Zimmer suggested that Lyft is "a more jovial and interactive social experience than simply getting into the back of a slick black car" (Olanoff, 2012). Additionally, Lyft's early business model rested upon what the company called "donations." After each ride, riders would be asked to "donate" to their drivers. Lyft's app would have a "suggested donation" amount calculated based on the location and length of the trip, but riders could adjust their "donations." Lyft drivers received 80% of the donation, whereas Lyft took the remaining 20%. The donation-based model attempted to evade taxi regulations. Nonetheless, Lyft reportedly sent a warning message to riders who paid less than the suggested amount—"low donations may make it hard for you to get rides in the future" (ABC7, 2013). Lyft drivers could also reject riders' requests based on the latter's payment history. As a response to public criticism, the company eventually replaced the donation-based model with a mandatory pricing model in 2014.

Figure 2. Screenshot of Lyft's Website on January 26, 2015 (Source: The WayBack Machine)



Lyft gained traction and became the main business of Zimride. In May 2013, the company renamed itself Lyft, and later sold its carpooling business to Enterprise Rent-A-Car. Similar to Uber, Lyft incorporated a two-way 5-star rating system at the beginning of its business. Passenger ratings, explained Lyft, were a means to protect drivers: "Drivers rate passengers after every ride, and you'll be able to see their rating before accepting a pickup. Every passenger also signs up for Lyft using Facebook and provides valid credit card information for identity purposes." Passengers, meanwhile, could rate their drivers and their ratings mattered in two ways. First, if a passenger rated a driver below 4 stars, they would never be matched on the platform. Second, Lyft would deactivate drivers with an average rating lower than 4.5 stars. The rating system, as Lyft advertised on its website, was "all about community."

Despite Lyft's emphasis on community, Lyft and Uber shared similarities in terms of their business models. For example, both Uber and Lyft claim to target drivers who want to be their own boss and enjoy flexibility. Drivers are classified as independent contractors. Lyft developed "Lyft Line" (or "Shared rides"; c.f. UberPool) and "Lyft Lux" (c.f. Uber Lux). In 2016, Lyft replaced its iconic pink mustaches with a new device called "Amp"—"a two-way beacon the company claims

will help riders and drivers find each other faster and more easily than ever—after Uber tested a similar color-matching system (Stinson, 2016). As Uber has implemented its rewards program for drivers since 2018, Lyft accordingly expands its rewards program to drivers in 2020. Additionally, Lyft and Uber have offered bonuses (e.g., Lyft "Prime Time" versus Uber's "surge") to attract and retain drivers.

Yet, there were a few key differences between Lyft and Uber in the past. First, Lyft has allowed riders to tip their drivers on the app since 2012, whereas Uber had long opposed the idea of tipping (Uber Under the Hood, 2016) and only incorporated this function as part of its "180 Days of Change" campaign in 2017. Lyft (2018) further changed its app interface to incorporate a default tipping option in 2018. The suggested tips were calculated based on the percentage of the total fare, rather than a preset tip amount (e.g., "$2"). Second, Lyft (2019) has emphasized its "driver-centric" approach by offering 24/7 phone support for drivers and establishing in-person local Hubs. In 2016, Lyft also established a "Driver Advisory Council" to invite drivers who exemplified its community guidelines to share their opinion with Lyft. In fact, Lyft (2018) referenced the Council when incorporating the default tipping option. Third, since 2018, if a rider does not rate a Lyft driver, the latter's rating is by default a 5-star rating. As a result, it seems that drivers had a higher level of satisfaction with Lyft than Uber, according to a non-probability sampling survey conducted by "The Rideshare Guy" in 2018.

Lyft has become a public company since 2019. Compared to Uber, Lyft's (2019) future business plan—what it calls "multimodal platform"—remains largely focused on transportation. Lyft's multimodal platform includes ride-hailing services,

bikes and scooters, public transit, and autonomous vehicles. Its ultimate goal is to "simplify travel decision-making and expand the potential uses for our platform, such as subscription plans, commuter services, first-mile and last-mile services and university safe rides programs" (Lyft, 2019, p. 9).

Algorithmic Labor Process and Metrics on Ride-hailing Platforms

Uber's and Lyft's labor processes are largely similar. Drivers needs to log onto the worker-facing app and wait for a ride request. Usually, the interviewees waited for their first ride either in their house or an area in which they anticipated a high demand of ride requests. Drivers can see a colored map (Uber's surge map or Lyft's "Personal Power Zones") that claims to visualize areas with a high demand of ride requests and more earning opportunities. The two platforms also offer other on-demand bonuses such as Uber's "Quest" and Lyft's "Ride Streaks." Once drivers receive a ride request, they can see limited information about the trip, including the rider's name, fare upfront, and pickup estimated time arrival (ETA). Depending on drivers' working location and rewards program status (Uber Pro or Lyft's rewards program), drivers may have access to the trip duration, pickup location, and trip time. After the passing of AB5 in California, both companies made available the details about trip information to drivers in California. As Uber (2019b) said, such information could help these drivers to "earn on their own terms" and "make informed decisions." Drivers have 15 seconds to accept the request. Then, drivers commute to the pickup location and navigate to the destination. During the ride, drivers can rely on the platform's navigation system or other navigation apps. After the ride, the driver and rider can rate each other's performance on a 5-star scale. The driver must rate their rider *before*

accepting another ride, whereas the rider can decide whether they want to rate the driver. Additionally, both Uber and Lyft charge service fees on each ride (excluding customer tips), but the amount varies per trip.

Table 4 in the Appendix offers a summary of (a) the kinds of worker-related metrics that the platforms currently incorporate into its platform; (b) methods of calculating customer ratings; and (c) dispute process. There are three main differences between Uber's and Lyft's design of their metrics. First, Uber calculates drivers' average ratings based on their last 500 rated trips, while Lyft only accounts for the last 100 trips. Second, Lyft automatically removes the single lowest rating from drivers' account for every 100 rides. Third, if riders do not rate their drivers within two hours, Lyft drivers will automatically receive a 5-star rating.

Both companies have used 5-star rating systems since their inception. In 2014, Uber sent a guide to the drivers to explain the methods of calculating ratings, the goals of using the rating system, and strategies for getting a good rating. The official guide stated, "Your [drivers'] average rating is the first thing that Uber uses when evaluating your [drivers'] partnership and the system" (Uber, 2014, p. 2). Uber may deactivate the "lowest-quality drivers" to protect "the quality of the Uber system as a whole" and ensure that riders can enjoy an efficient and safe ride-hailing service (p. 3). Accordingly, Uber encourages drivers to offer additional amenities (e.g., water, snacks, and cell phone chargers), keep their vehicle clean, and dress professionally. Lyft, by contrast, generally suggests drivers to consider ways to improve their performance if their average ratings fall below 4.8.

The improvement of user experience has been a commonly used rhetorical

justification for Uber and Lyft to change the design of their rating system. In 2015, for example, both companies allowed riders to check their ratings to increase the "transparency" of the system. Lyft also notified riders when drivers gave them a 5-star rating (Soper, 2015). Yet, Lyft drivers were used to receive a copy of all customer feedback (Ince, 2017), but now they only receive Lyft's weekly feedback summary (Avedian, 2019). After Uber's "180 Days of Change" in 2017 and Lyft's implementation of ratings protection in 2018, their rating systems became very similar. Both companies require riders to briefly explain why they would give a rating below 5 stars. Nonetheless, the two companies do not disclose much information about their dispute process. Uber also explicitly mentions that no individual ratings can be removed upon drivers' or riders' requests.

Both companies have offered rewards programs for drivers. The rewards programs can be considered a form of gamification: "following rules, accumulating points, obtaining prizes" (Turow, 2017, p. 167). In Turow's (2017) analysis of loyalty programs for customers, he found that retailers might consider such programs as means to establish emotional connections with "high-value" customers and incentivize them to stay tuned for the stores. Through loyalty programs, retailers can collect and track customers' personal data. Within the context of this project, gig workers have already given up personal data to labor platforms when they work on the platforms. But the rewards programs may still be useful for retaining workers on the platform. A worker's reward status is dependent on how many rides they offer and how well they perform in accordance with platform-based evaluative standard. In this vein, platform owners can use rewards programs to incentivize workers to stay active on the platform

and incite them to meet the needs of their customers. The programs, meanwhile, can set normative standards about "good" workers who may receive better treatment on the platform.

Uber has begun its Uber Pro program in 2018, while Lyft has extended the scope of its rewards program to cover drivers. Uber and Lyft drivers are automatically enrolled into the rewards program. Each program period lasts for three months. For Uber drivers, they begin at the "Partner" tier and collect points to ascend to the "Gold," "Platinum", and "Diamond" tiers.[14] For each ride, drivers can earn one point for every trip they complete, but they may earn extra points when they drive in designated times and areas. At the Diamon tier, drivers can "unlock" the premiere level of benefits such as "100% tuition coverage" through Arizona State University's online platform, priority status for airport pickups and preferable reimbursement on fuel charges. All Uber Pro points will be reset after the three-month period. To maintain the status, drivers must provide "quality service," represented by a customer rating of 4.85 or above, a cancellation rate of no more than 4%, and an acceptance rate of at least 85%.[15]

Lyft's rewards program, similarly, has three tiers: "Silver," "Gold," and "Platinum." Lyft drivers can earn one Qualifying point for each dollar they earn during busy hours, but the eligible earnings do not include cancellation fees, ride

[14] Uber sets different points thresholds in different cities and may modify them in different program periods. For example, a Chicago driver may need to get 600 points to unlock the "Gold" tier, 1200 points to unlock the "Platinum" tier, and 1800 points to unlock the "Diamond" tier. A driver in Denver, by contrast, may need to get 300 points for "Gold," 600 points for "Platinum," and 1200 points for "Diamond."

[15] Uber has removed the acceptance rate requirement in California as a way to demonstrate that the company is not controlling the workers.

bonuses, and other bonuses. The Qualifying points will be reset at the end of each program period. In addition to unlocking tiers, Qualifying points can also be redeemed for ride credits ("Silver" tier) and cash ("Gold" and "Platinum" tiers). In this sense, the rewards program can provide a monetary incentive for drivers to continue working on the platform. Yet, having a high Qualifying point is only one of the criteria for unlocking tiers. Lyft, meanwhile, calculates drivers' driving score, based on their ratings, acceptance rates, and cancellation rates in their last 100 rides. Lyft does not disclose the exact formula for calculating one's driving score. Drivers can look at their score on the rewards program screen of Lyft's app. This driving score is only relevant to the rewards program. Drivers must maintain at least 60% of driving score to unlock "Silver" and 80% of driving score to unlock the other tiers. In order to unlock the tiers of "Gold" and "Platinum," drivers need to have a rating higher than 4.90 and a 90% acceptance rate. The acceptance rate is required in all cities, except for California.

Considering algorithmic labor control through the lens of direction, evaluation, and discipline (Kellogg et al., 2020), metrics as a form of evaluation influences algorithmic direction and discipline in at least three ways. First, ratings, cancellation rates, acceptance rates, and rewards points matter for Uber and Lyft drivers to maintain and unlock their tier. Here, the rewards programs can be seen as a form of algorithmic discipline by offering "rewards" for drivers. One of the important "rewards" for Uber and Lyft drivers is the increased visibility of trip information (e.g., pickup location and trip time). This leads to the second point about the direction of work tasks. Drivers could possibly make use of the details about trip information to organize their work schedule and spatial movement. But with the implementation of

the rewards programs, such information becomes only available to drivers who continuously provide ride-hailing services and receive a high rating from their riders Third, metrics, particularly ratings can affect drivers' continuity of employment on the platforms. If drivers' average ratings fall below the location-specific minimum threshold, they are at risk of being deactivated.

TaskRabbit

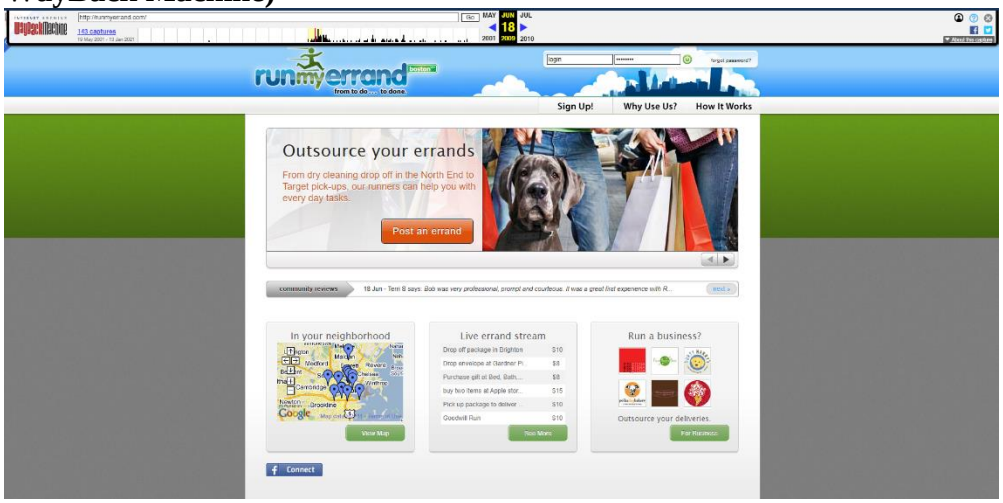
Among the platforms I studied, TaskRabbit has significantly altered its organization of work tasks throughout its history—from a bidding marketplace system to an on-demand matching system. TaskRabbit is also different from the other platforms in two ways. First, it allows Taskers to set their own hourly rates, which according to Stacy Brown-Philpot—the former CEO of TaskRabbit—contributes to the building of "relationships and community bonds" on TaskRabbit (Wired, 2018). Second, Tasker-related metrics can determine the visibility of their online profiles.

A Brief Overview of TaskRabbit's Development Trajectory

TaskRabbit was formerly named "RunMyErrand" between 2008 and 2010. As the founder Leah Busque envisioned, the company would have "incredible underlying power" to build "a sense of community" (Truong, 2011). The core idea of the company is that "neighbors would come to your rescue – they would do your daily errands" (Scholz, 2017, p. 47). From the company's perspective, daily errands were essentially the tasks that people "don't have the time, ability, or desire to do themselves." Using its "errand sharing model," the company promised to "save time," "save money," "contribute to green living," and "support your local community." At that time, RunMyErran was a web-based and mobile marketplace (see Figure 3-1) that

connected “senders” (i.e., service recipients) with “runners” (service providers).

Figure 3-1. Screenshot of RunMyErrad’s Website on June 18, 2009 (Source: The WayBack Machine)

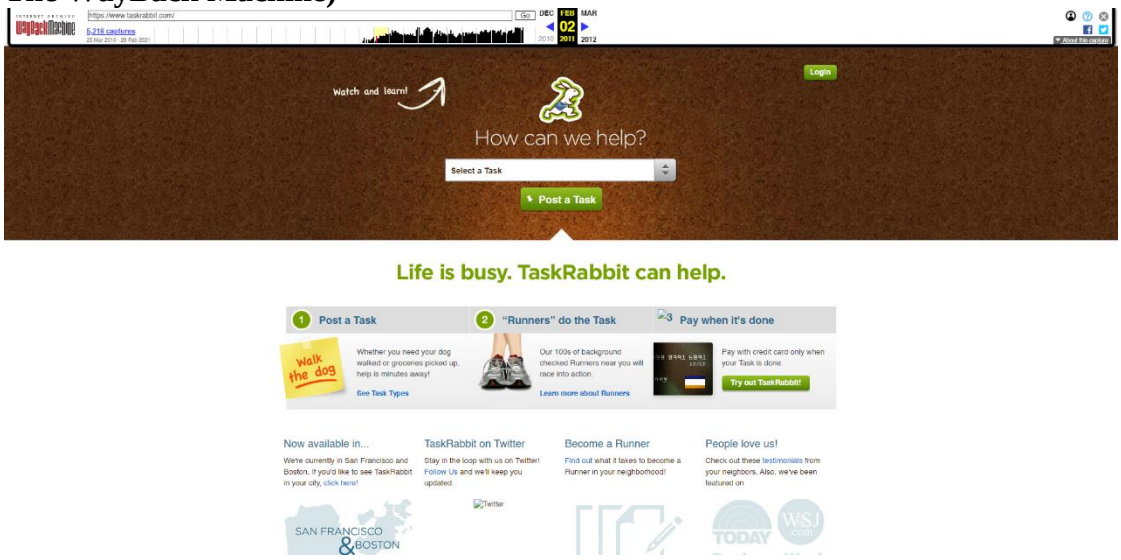


Specifically, senders could post any small tasks and set the rates on the platform. Once a task was posted, a group of “prequalified” runners who passed the company’s background check would receive an alert via e-mail or text messages. Runners could then accept the task or counteroffer the price. To inspire senders, RunMyErran created a list of “top errand ideas” such as “wait in line for sports tickets,” “address and mail holiday cards,” and “pick up a 40-lb bag of dog food.” On the website, people could see lists of runner profiles with star ratings (5-star scale) and reviews, available tasks (“running now”) and recently completed tasks. The list of recently completed tasks included the content and price of the tasks as well as customer ratings and reviews. The ratings and reviews, or what Leah Busque called “reputation engine” distinguished the company from Craigslist (Truong, 2011).

As the company grew and received over $1 million investment from venture capital funds in 2009 and 2010, Leah Busque decided to rebrand the company as TaskRabbit in April 2010. With this rebranding campaign, TaskRabbit launched a

series of changes in its platform (see Figure 3-2 for a screenshot of TaskRabbit's website on February 2, 2011). First, it began focusing on a predefined category of tasks beyond errands. In 2011, the categories included shopping, donation pickup, misc. pickup/drop-off, moving help, house chores, and something else (e.g., restaurant pickup, writing, and computer help).

Figure 3-2. Screenshot of TaskRabbit's Website on February 2, 2011 (Source: The WayBack Machine)



Second, TaskRabbit reframed "senders" as "TaskPosters" and "runners" as "TaskRabbits" (and more recently, "Taskers"). The company also considered runners or TaskRabbits as "entrepreneurs" who were "in charge of their own schedules and work environments" (The Hutch by TaskRabbit, 2010). Third, TaskRabbit launched its iPhone app in July 2011. The app enabled service recipients and providers to find the tasks based on location and provided them with suggested pricing for tasks (The Hutch by TaskRabbit, 2011). TaskRabbit, meanwhile, remained using its bidding system. TaskPosters could assign a price for their task or leave the task open for bids. If they chose the latter option, the system would allow them to take the lowest bid. At

this time, TaskRabbit's transaction fee was between 12 and 30 percent of workers' wage (Scholz, 2017).

Of particular relevance to this project, there were two major changes in TaskRabbit in 2014. First, TaskRabbit has introduced a rewards program—"TaskRabbit Elite program"—for highly active workers with at least a 4.9 star rating (The Hutch by TaskRabbit, 2014a). This program does not offer monetary rewards; instead, it recognizes workers via "the special badge on their profile and offers." Second, TaskRabbit has shifted from a bidding system to a task matching algorithm (The Hutch by TaskRabbit, 2014b). While TaskRabbit claimed to introduce this change "based on months of user research" (The Hutch by TaskRabbit, 2014b), the company reportedly adopted this model because of the decline in bids (Newton, 2014). Specifically, TaskRabbit's matching algorithm would show clients a list of suitable Taskers, based on the latter's availability and past behavior (e.g., ratings, experience, and hourly rates) on the platform. TaskRabbit also requires all Taskers to set their hourly rates in advance, so as to reduce "pricing anxiety" (The Hutch by TaskRabbit, 2014b). Taskers are no longer allowed to view all the available tasks.

In April 2016, Stacy Brown-Philpot became the new CEO of TaskRabbit, as Leah Busque had decided to step down.[16] Between 2016 and 2017, Stacy Brown-Philpot significantly expanded TaskRabbit's presence from 19 cities to 40 cities (Wolfe, 2017). She also oversaw the sale of TaskRabbit to IKEA in 2017 and described this action as "opportunistic" to expand TaskRabbit to be a "larger

[16] This is the second time Leah Busque stepped down as TaskRabbit's CEO. In October 2011, she decided to focus on product development and hired former Hotwire CEO Eric Grosse to work as TaskRabbit's CEO. However, she soon resumed the CEO in June 2012.

platform" (Brown, 2017). After IKEA's acquisition, she remained the CEO until August 2020. TaskRabbit is still an independent company, while partnering with IKEA to offer furniture assembly services.

Since August 2020, Ania Smith—the former head of the Courier Operations at Uber Eats—has become TaskRabbit's CEO. In 2020, the company has introduced 12 categories of tasks such as smart home installation, room measurement, and snow removal (The Hutch by TaskRabbit, 2020). Due to the COVID-19 pandemic, the company has also introduced contactless and virtual tasks. Yet, the impacts of the pandemic on the company's development remain unclear.

TaskRabbit's Algorithmic Labor Process and Metrics

Among the Taskers I interviewed, Kooper and Zoey were the only interviewees who had experienced TaskRabbit's bidding system. Both of them had positive attitudes towards TaskRabbit's change. Zoey described the previous bidding process as a "horrible" experience because many people would apply for a task and the client would only take the lowest offer. Kooper, similarly, considered the change as "a win-win because … you can also set your prices at a higher threshold if you have more experience."

That said, the rest of the interviewees had never used the bidding system. Currently, the first step for working on TaskRabbit is to set up an online profile. Taskers' online profile shall include their personal information and specify the skills, schedule, and hourly rates. Taskers can set different hourly rates for different categories of tasks. Taskers can also decide whether they would accept same day requests. Some of my interviewees also mentioned the tools they had for home repairs,

mounting, and furniture assembly on their profiles. However, it is noteworthy that all of the interviewees told me that they did not update their profiles frequently, except for their weekly schedule and hourly rates. On the clients' side, they need to input their address and briefly describe their tasks. Then they will be able to see a selected list of Taskers sorted by TaskRabbit's algorithm. As previously discussed, TaskRabbit's algorithm accounts for Taskers' availability, experience, and ratings. Clients can also select the options to narrow the search results to a list of "Elite Tasker" and "Great Value" (i.e., Taskers with a high rating and a low rate). They can view Taskers' profiles and ratings before selecting a specific Tasker.

After the client selects a Tasker, the Tasker will receive a notification on the worker-facing app. Before accepting the task, the Tasker needs to communicate with the client to confirm the details about the task, using the "chat" function. After that, if the Tasker decides to accept the task, s/he needs to confirm the time and estimate the length of the task. There is a one-hour minimum for all tasks. After completing the task, the Tasker need to create an invoice on the app to confirm the hours s/he worked and specify any additional expenses for completing the task. The Tasker can also leave a closing message for the client. After the Task is invoiced, the client can leave a rating and a public written review for the Tasker. Unlike Uber and Lyft, TaskRabbit charges a 15% service fee of the total price on clients rather than Taskers.

There are four kinds of worker-related metrics on TaskRabbit, including 5-star customer ratings, response rates, acceptance rates, and reliability rates (see Table 4). In the early days of RunMyErrand and TaskRabbit, service recipients and providers could rate one another's performance on a 5-star scale. Yet, as TaskRabbit has

transitioned to the current task matching system, Taskers are no longer able to rate their clients. Taskers' 5-star ratings are categorized as "positive" (3 to 5 stars) and "negative" rating (1 to 2 stars). Yet, prospective clients can look at all individual ratings and reviews. Taskers can dispute an unfair rating to the company if they provide valid documentation. Additionally, the other three types of metrics are expressed as percentages. If Taskers accept all the request and complete those tasks, they have perfect metrics (i.e., 100%) in these three categories. TaskRabbit suggests that Taskers should maintain at least 75% acceptance rate, but they do not deactivate any Taskers based on metrics. While TaskRabbit renders all kinds of metrics visible to workers and customers, the company does not clearly explain how such metrics influence the matching algorithm. Put another way, Taskers know that their metrics are one of the key factors that influence the visibility of their online profiles but do not understand how TaskRabbit actually uses their metrics in the matching process.

Metrics are connected with the visibility of Tasker profiles on the platform in two ways. First, Taskers' metrics can affect their standings in search results, but they do not affect what work-related information Taskers can see before accepting a task. Second, Taskers' metrics, particularly ratings, are basic requirements for qualifying the platform's "Elite Taskers" program. Taskers need to maintain at least 98% of positive ratings and complete at least 10 tasks in the previous month to participate in this program. This program provides "Elite Taskers" with a badge on the platform, a higher suggested hourly rate, priority support, and access to the company-led "Tasker Elites" Facebook group. Yet, even if Taskers lose their "elite" status, they will not be removed from the Facebook group.

Delivery Platforms: DoorDash, Uber Eats, and Instacart

A Brief Overview of DoorDash's Development Trajectory

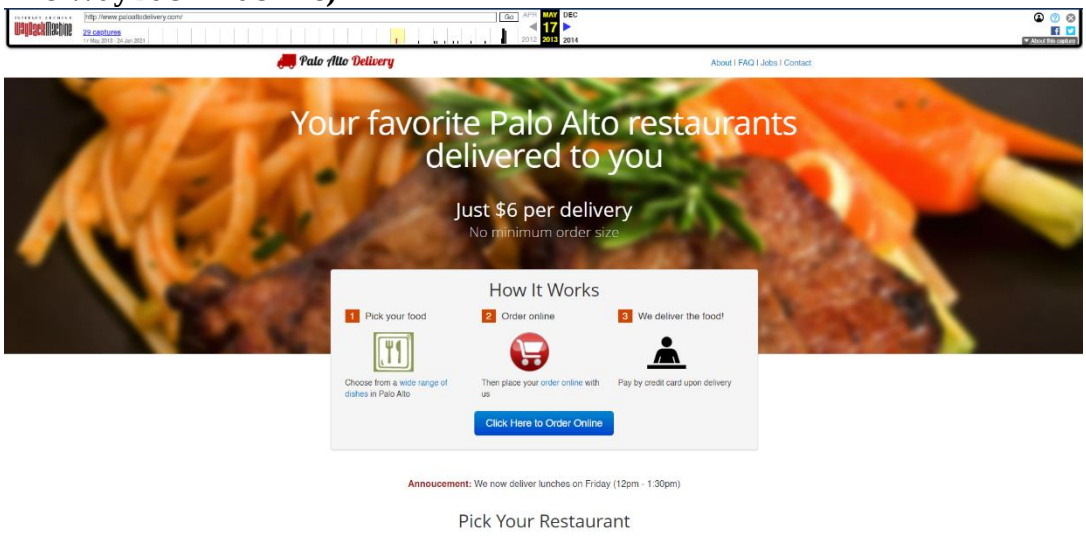
Originally named as "Palo Alto Delivery," DoorDash was found by four Stanford students—Tony Xu, Stanley Tang, Andy Fang, and Evan Moore— in 2013 (Luce, 2020). As Tony Xu and Evan Moore recalled, the four co-founders started brainstorming the idea in Fall 2012, because "most restaurants in the area [i.e., downtown Palo Alto] didn't deliver" (Luce, 2020). The four co-founders, accordingly, conducted interviews with small business owners in the Bay Area and found that delivery infrastructure was unavailable for these owners (How to Start a Startup, 2014).

In January 2013, they decided to experiment their delivery service by launching Palo Alto Delivery's website (see Figure 4). As Stanley Tang reflected in 2014, "We didn't have any drivers; we didn't have any algorithms; we didn't have a backend; we didn't spend six months building a fancy dispatch system" (How to Start a Startup, 2014). They found the PDF menus of the restaurants and uploaded them on the website. The four co-founders were the delivery drivers. Six months after founding the business, they completed 217 deliveries and decided to apply for the seed funding from Y Combinator—a seed money accelerator program for start-ups[17] (Buchheit, 2020). They eventually received $120,000 and scaled up DoorDash to offer delivery services "in Palo Alto and Mountain View in an average time of 45 minutes" on its

[17] Y Combinator's program is essentially about funding start-ups (in exchange for future equity), offering advices to the start-ups, and providing networking opportunities. In addition to DoorDash, Instacart, Airbnb, and Dropbox are some of the companies that participated in the program.

website or via a mobile app (Gallagher, 2013). DoorDash, meanwhile, developed three separate apps for its customers, drivers, and restaurants.

Figure 4. Screenshot of Palo Alto Delivery's Website on May 17, 2013 (Source: The WayBack Machine)



DoorDash has positioned itself as a technology and logistics software company since 2013. The company advertised delivery drivers, or what they called "Dashers" as either part-time or full-time independent contractor positions, while promising that the company "promoted drivers in the past to other positions within DoorDash." It also required drivers to have an iPhone, car, insurance, clean driving record, and friendly attitude. DoorDash, nonetheless, was sued over its misclassification of worker status, in part because the company required Dashers to red T-shirts with DoorDash logo for delivering order and relied upon independent contractors as its major workforce (S. Levy, 2015). The company eventually paid $3.5 million to settle class-action lawsuit and updated its deactivation policy to clarify the conditions under which Dashers

might be deactivated in 2017 (DoorDash, 2017a)[18]. In the deactivation policy, we can see that maintaining at least 4.2 customer rating (out of 5-stars) and 80% completion rate is essential for Dashers to stay on the platform. Yet, DoorDash has framed that such metrics are crucial for ensuring the success of the platform and maintaining a high-quality customer service. The deactivation policy still insists on the company's classification of independent contractors.

There are two notable changes in DoorDash's rewards and metrics for delivery drivers. First, DoorDash tested a pilot monthly rewards program—the "Top Dasher Program"—for delivery drivers in 2019. It required delivery drivers to have at least 4.70 customer ratings, 95% completion rate, 100 completed deliveries, and 200 life-time deliveries by the last day of each month. A "top Dasher" can dash anytime and receive more deliveries. Notably, DoorDash requires delivery drivers to schedule their working time in advance, particularly in a market where the demand for delivery services is low. However, the scheduling did not seem to be relevant to my interviewees, possibly because the COVID-19 pandemic has significantly increased the demand for DoorDash's service. The rewards program has currently been implemented in the United States. Additionally, DoorDash (2020b) has attempted to "make ratings more actionable" since December 17, 2020. Accordingly, the worker-facing app allows delivery drivers to "see how many ratings have been excluded because of factors outside of your control, like longer wait times at restaurants." The app also provides more detailed written feedback from customers and customer

[18] This is only one of the class-action lawsuits against DoorDash. DoorDash (2020a) listed at least 18 lawsuits against the company and Caviar (another delivery platform acquired by DoorDash in 2019) between 2017 and 2020.

compliments (i.e., "Quotes from customers") as well as "tips from expert Dashers" on how to strengthen the ratings. However, DoorDash's deactivation policy remains unchanged. It is also noteworthy that the changes in ratings have been implemented *after* I conducted the interviews with DoorDash delivery drivers.

A Brief Overview of Uber Eats' Development Trajectory

Uber launched Uber Eats (formerly named "UberFresh") in August 2014. In 2014, "UberFresh" was built into Uber's app—much like other ride options—and only available in San Monica and Los Angeles. Uber renamed the service as Uber Eats in April, and launched a separate app for its delivery service in August 2015 (Tepper, 2015). At the time, Uber Eats focused on lunch and dinner options in a set time period. It charged a flat delivery fee ($3 in all cities except New York; $4 in New York) (Tiku, 2015). But the company has decided to change its pricing model to charge a delivery fee based on distances in 2018 (Lee, 2018). Currently, Uber Eats charges their customers a distance-based delivery fee and a 15% service fee of the order (between $2.5 and $4). It also has a one-time activation fee ($350) and marketplace fees (from 15% to 30% of each order) for restaurants. It takes about 25% service fee from drivers' earnings (excluding customer tips).

Indeed, Uber Eats was initially an experiment of what Uber called "Uber Everything"—"a startup within a startup" (Goldin, 2016). The company envisioned that Uber Eats would become "the fastest, easiest way to start earning," because delivery drivers do not even need a vehicle. The delivery drivers "can deliver with a bicycle, a scooter, or even deliver on foot" depending on where they work. In March 2016, Uber Eats experimented a new option called "instant" delivery in New York

City, promising to delivery pre-set lunch options within 10 minutes from the time of ordering, but this option was abandoned four weeks later (Griswold, 2016). As the platform developed, Uber used an aggressive approach to expand its food delivery service. In 2017, Uber Eats has become available for 24/7 in major cities such as New York (Mascali, 2017). Uber Eats also reportedly only "operate[s] in markets where it could be one of the top two delivery services, and exit all the others" (Abril, 2020). Uber Eats has become the faster growing part of Uber's business.

Despite such growth, the design of Uber Eats' app remains largely the same over the course of three years.[19] Customer ratings, acceptance rates, and cancellation rates have been adopted on the platform since 2015. However, Uber does not use a 5-star rating system on Uber Eats; instead, it uses a "thumbs up/ thumbs down" rating system. The Uber Pro program I discussed earlier also applies to Uber Eats delivery drivers.

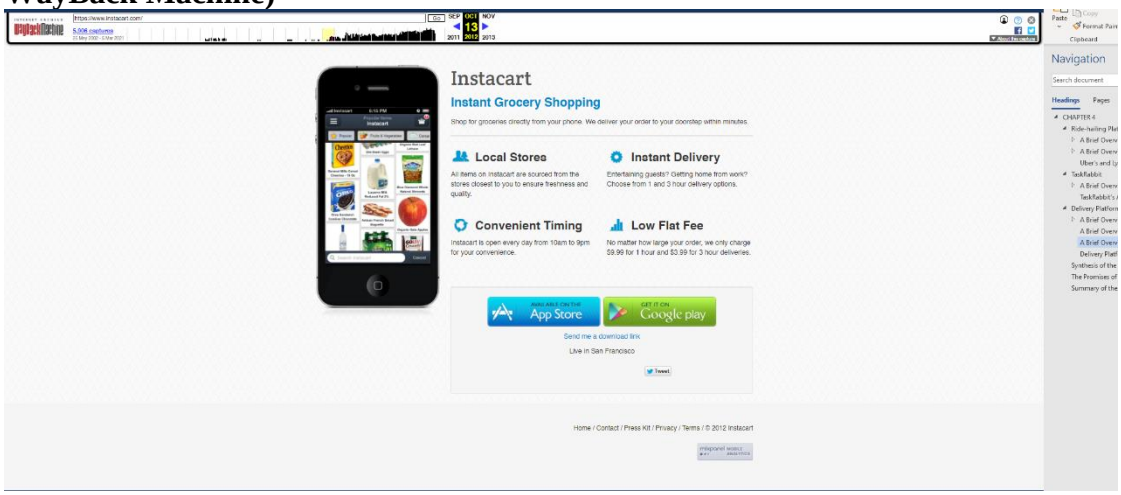
A Brief Overview of Instacart's Development Trajectory

Apoorva Mehta, a former Amazon engineer, found Instacart in San Francisco in May 2012. Its main service was to "allow customers to order groceries from their favorites stores and have them to deliver to their door within one hour" (Y Combinator, 2018). The company charged a flat delivery fee ($3.99 for any orders above $35. Akin to the other labor platforms I studied, Instacart does *not* have its own grocery delivery infrastructure. Consider the counterexamples of Amazon Fresh and

[19] In September 2019, Uber's "main" ride-hailing app has enabled drivers and customers to use Uber Eats in order to become "the operating system for your everyday life" (Hawkins, 2019b). However, drivers and customers can remain using the Uber Eats app. The labor processes of Uber and Uber Eats remain unchanged.

Walmart grocery delivery services which were Instacart's main competitors. Both Amazon and Walmart have warehouses for selling and storing groceries; however, Instacart directs what it calls "personal shoppers" to buy groceries for their customers via its mobile app (see Figure 5 for the company's website in 2012). As Mehta stated, "Our goal is to offer an Amazon-like experience without building any the infrastructure, using crowdsourcing" (Kirsner, 2013). Instead, Instacart provided shoppers with flexible employment opportunities. Hence, Mehta suggested that Instacart "can launch new cities within days," as opposed to the warehouse model (Frank, 2013).

Figure 5. Screenshot of Instacart's Website on October 13, 2012 (Source: The WayBack Machine)



While Instacart emphasized its ability to allow customers to order from any local grocery stores (Frank, 2013), the company—like DoorDash—initially did not have formal partnership with many stores it listed on the app. One such example is Trader Joe's which was listed as one of Instacart's offerings between 2012 and 2013. As Mehta recalled in a Y Combinator (2014)'s talk,

There is no API or website which had the item catalog for Trader Joe's, so the

> only way we could actually get all of the item information was to buy one of every single item at Trader Joe's, take it to a studio, take pictures of all those things, and then put them into our catalog.

In October 2013, Trader Joe's required Instacart to remove it from the offering until a formal agreement could be reached (Frank, 2013). In San Francisco, similarly, Instacart was denounced for its lack of formal agreements with grocery stores (Somerville, 2014). This led Instacart to focus on building partnerships with retailers (Narula, 2015). Yet, such partnerships are not permanent, which might threaten Instacart in-store shoppers' employment opportunities. As I described in Chapter 3, in-store shoppers are Instacart's part-time employees and mostly work in a specific grocery store. After Amazon acquired Whole Foods Market, the grocer subsequently announced to stop partnering with Instacart in December 2018. At that time, there were 1,415 in-store shoppers at Whole Foods Market (Instacart, 2018). As a result, over 200 in-store shoppers would have lost their jobs in February 2019, unless they were willing to work at a different store in their area.[20]

The hiring of in-store shoppers also distinguishes Instacart from the other labor platforms I studied. Instacart is the only platform that recruit both part-time employees (i.e., in-store shoppers) and independent contractors (i.e., full-service shoppers). Before June 2015, all of the shoppers were full-service shoppers. Instacart provided full-service shoppers in Chicago and Boston with an option to become part-time

[20] As Instacart had to close all of their in-store shopping operations at Whole Foods Market, 1,415 in-store shoppers would be at risk of losing their jobs. The first phase of the transition impacted 243 in-store shoppers. Instacart (2018) also expected to place over 75% of the impacted shoppers in another store. Yet, it is unclear how many shoppers eventually lost their jobs.

employees, possibly because there was a class-action lawsuit against the company's misclassification of worker status (Alba, 2015). Currently, in-store shoppers only account for a small portion of Instacart's workforce. Their employment status, however, enables them to unionize themselves. In February 2020, ten Instacart in-store shoppers at a grocery store in a suburb of Chicago voted to form a union (Gurley, 2020). Yet, Instacart fired nearly 2,000 in-store shoppers—including the shoppers who voted to unionize—in January 2021 to accommodate the growing demand for grocery pickup services, as the company stated (Schiffer, 2021). As Instacart expands its workforce of full-service shoppers during the pandemic, in-store shoppers are at risk of being fired.

In addition, Instacart has been criticized for its pricing model. Here, I focus on full-service shoppers' earnings, rather than in-store shoppers who receive a flat hourly rate, or customers.[21] Initially, shoppers earned a 40-cent commission per item and 100% of customer tips (Lieber, 2019). The app also suggested a 10% customer tip per order (Glaser, 2019). In September 2016, Instacart (2016) replaced the customer tips with an optional "service fee." The main difference was that Instacart—rather than the shopper—would collect all of the service fees. As Del Rey (2017) reported, "The startup's explanation was that Instacart workers were too reliant on tips — around 80 percent of orders had one — and that the service fee would allow the startup to pay everyone a more reliable wage." Due to the strong opposition from shoppers, Instacart

[21] The main criticism of Instacart's pricing model for customers is the ambiguous makeup percentage. It was until 2015, Instacart showed the price difference between the app and stores. The markup percentage is estimated to be about 15% to 23%, but it also varies across stores and products (Deng, 2016). Yet, Instacart reportedly would deactivate shoppers who give the store receipt to their customers (Lieber, 2019)

added back the tipping feature, but put the feature under the option of service fees and made it more for customers to find the feature (Del Rey, 2017). In 2018, Instacart also changed the customer suggested tip from 10% to 5% per order (Glaser, 2019).

Starting from 2018, Instacart has used an algorithmic pricing system to replace the piece-rate commission system (Griesbach et al. 2019). The baseline pay of each full-service batch was between $7 and $10 (Lieber, 2019). Instacart also provided shoppers with a $3 "quality" bonus for every 5-star rating they received (Karls, 2019). However, Instacart does not specify the factors that determine the final earnings (Griesbach et al. 2019; Lieber, 2019). In January 2019, a group of full-service shoppers denounced Instacart for its "black-box pay algorithms" (Working Washington, 2019). As they complained, the deployment of the algorithmic pricing model did not provide a clear breakdown of the earnings and encouraged shoppers to accept low-paying batches. They also noted that "some of us have seen wages lowered by 30–40% overall" and demanded for "a predictable, transparent pay structure."

As part of the change in the pricing system, Instacart has also hidden the delivery address *before* full-service shoppers accept a batch to protect customer privacy (Captain, 2019). Yet, shoppers found it more difficult to determine which batch they should accept due to the limited information about the delivery order (Captain, 2019). The changes in the algorithmic pricing system and tipping feature eventually led thousands of shoppers to go on a 3-day strike between November 3rd and 5th in 2019 (Glaser, 2019). Two days after the strike, Instacart cut the "quality bonus" because it "did not meaningfully improve quality," as the company claimed (Gurley, 2019). Instead, Instacart decided to implement other kinds of bonuses that

encouraged shoppers to work in “busy areas” and during “busy hours” (Instacart, 2020b). Full-service shoppers’ ratings affect their access to batches and continuity of employment on the platform. As the COVID-19 pandemic has significantly increased the demand for delivery services, Instacart is planning to become a publicly traded company.

Delivery Platforms’ Algorithmic Labor Processes and Metrics

The labor processes on delivery platforms are similar to that of ride-hailing platforms. Delivery drivers log onto the worker-facing app and wait for delivery request. DoorDash and Instacart delivery drivers may need to schedule their working hours on the app in advance. When they receive a request, they can see the minimum amount of earnings, customer tips, and restaurant address. Then, delivery drivers have a limited time (e.g., 15-30 seconds on Uber Eats and 45 seconds on DoorDash) to accept an order. Then they need to go to the restaurant to pick up the order. After that, they commute to the customer’s address and deliver the food. Due to the COVID-19 pandemic, delivery drivers and customers can request contactless delivery. In this case, delivery drivers follow customers’ delivery instructions to leave the food in a specific location. After that, delivery drivers take a photo to show where the food is placed and submit the photo to the app.

The labor process on Instacart is slightly different from that on DoorDash and Uber Eats because Instacart full-service shoppers can see a list of available batches, with the information about the order (e.g., what to shop and where the item is stored in the grocery store) and estimated earnings. A full-service shopper with a higher rating can see batches before other shoppers. Then, they compete with one another to pick

which batch they accept. If they do not want to accept any batches, they need to wait until the app generates another list of batches. Once they accept an order, they go to the selected grocery store to shop for the customers. If there are in-store shoppers in the grocery store, the full-service shoppers may skip the shopping process and directly pick up the order. Otherwise, they need to use the shopper card for checking out with the order. Like DoorDash and Uber Eats delivery drivers, Instacart shoppers can opt to conduct contactless delivery. After each order, customers can rate their delivery drivers on a 5-star scale (DoorDash and Instacart) or thumbs up/thumbs down scale (Uber Eats) (see Table 4 for each platform's method of calculation).

Platform-based delivery labor processes generally have three main characteristics. First, all of these platforms' algorithmic pricing models are opaque and arbitrary (Griesbach et al., 2019). Specifically, these platforms claim to account for a variety of factors—such as real-time demand and supply as well as delivery drivers' and customers' location—when calculating the rate for each order. Delivery drivers cannot anticipate what order they would receive and how much they could earn. DoorDash and Uber Eats drivers can only estimate the value of the order based on the minimum earnings. For instance, delivery drivers may receive "$3.75 to drive 30 miles" for completing an order (Gebel, 2020). Delivery drivers must evaluate the value of an order within a short period of time. While Instacart shoppers can look at a list of batches, they need to quickly accept a batch to compete with others.

Second, restaurants and grocery stores become important actors in the labor processes. The interviewees often mentioned "waiting" as an essential part of their work experience. It is not just about waiting for an order, but about waiting in lines for

picking up the order. Ideally, restaurants should begin preparing for the food after customers place the order. Therefore, delivery drivers can immediately pick up the food once they arrive at the restaurants. However, some restaurants may prioritize in-store customers than delivery orders (Gebel, 2020), or only start preparing for the food after delivery drivers' arrival. Delivery drivers then may need to spend additional time at the restaurants. Instacart full-service shoppers, like regular customers, wait in line for check-out. The platforms, however, do not compensate for the waiting time. While customer tips may partly compensate for the waiting time, it is noteworthy that customers are not required to tip their delivery drivers. As in the case of Instacart, we can also observe that platform owners can easily change their in-app tipping features and policies, which would immediately affect workers' earnings.

Third, metrics are connected with rewards and sanctions. Uber Eats Pro program is similar to the one for Uber drivers, but the program period is one month rather than three months. Uber Eats delivery drivers can earn one point per trip and have additional points during peak meal times (e.g., 3 points for lunch hours). As Uber Eats uses a thumbs up/thumbs down system, the rewards program requires drivers to maintain at least a 95% of satisfaction rating (i.e., thumbs up). DoorDash's Top Dasher program, similarly, requires drivers to complete at least 100 deliveries and maintain a minimum rating (4.7), acceptance rate (70%), and completion rate (95%) in each month. Instacart does not offer any rewards program, but shoppers' ratings are connected with their priority access to batches. What distinguishes these rewards from that on ride-hailing platforms is that the former does not influence the availability of work-related information. As for sanctions, all of these delivery platforms may

deactivate delivery drivers if their ratings are below the minimum threshold.

Synthesis of the Labor Platforms

To summarize, we can observe similarities and differences among these labor platforms' development trajectories and labor processes. First, all of these platforms have insisted on their classification of workers as independent contractors, despite facing class-action lawsuits. Additionally, Uber's and Lyft's operation might violate local transportation regulations. DoorDash and Instacart also offered delivery services without establishing formal partnerships with restaurants in their early days.

Second, these platforms changed their interfaces and labor processes over time. Yet, the changes were often about the implementation of algorithmic pricing systems, the display of work-related information on the worker-facing app, and the introduction of rewards programs. These changes can directly influence workers' earnings and their decision-making capabilities. The rewards programs, in theory, shall provide incentives to motivate workers to accept more orders and provide services that meet the needs of their customers. Such programs shall also recognize "good" workers who are active on the platform and have high ratings, which may create hierarchies among workers. When we look at platform-based metrics and rewards programs by the types of labor platforms, their design and rewards are similar, possibly because these platforms compete with one another.

Third, while algorithmic direction, evaluation, and discipline constitutes algorithmic labor control (Kellogg et al., 2020), these three aspects are interconnected. In the case of ride-hailing platforms, metrics become intertwined with algorithmic direction and discipline. On TaskRabbit, workers' metrics matter in affecting the

visibility of their online profiles. Finally, delivery drivers' metrics are mainly connected with algorithmic discipline. I will explore how such differences might shape power-relations between platform owners, workers, and customers in the next two chapters.

The Promises of Platform-Based Metrics

Much like anonymous online marketplaces (Dellarocas, 2003), workers and customers are total strangers in the gig economy. In order for platform owners to attract customers to use their service, they must build trust between workers and customers. Here, metrics seem to become technological solutions because they hold workers accountable. In this final section, I briefly examine how platform owners attempted to legitimize their use of metrics through the appeals to accountability and mutual care on a discursive level.

The companies claim that metrics, particularly customer ratings, can promote accountability. Here, the questions are accountability to whom and for what purposes? As in Uber's and Lyft's framing of their "two-way" rating systems, metrics can hold both workers and customers accountable because both parties can rate one another's performance. In the 2015 Federal Trade Commission workshop on the "sharing economy," Uber articulated its design rationale for the rating system:

> This rating system does three critical things: it (1) incentivizes high quality service, (2) establishes accountability, and (3) promotes courteous conduct, and helps to mitigate the discrimination that is all too common in traditional for-hire transportation (Uber, 2015, p. 5).

Uber considered that the "two-way" rating system can establish accountability because

riders "can expect highly rated drivers to provide polite and helpful service," whereas drivers "can feel comfortable picking up a highly rate [sic] rated rider" (p. 5). An underlying premise is that both drivers and riders may feel vulnerable because they are total strangers, especially at night or after drinking. Uber (2015) concluded that the rating system could promote "a safe and respectful work environment" (p. 5). As Lyft (2019) similarly put this, the rating system "helps promote the safety and comfort of the Lyft community by offering a channel for drivers and riders to provide instant feedback on their Lyft experiences" (p. 148).

By contrast, TaskRabbit, DoorDash, and Instacart do not allow workers to rate their customers. Yet, they similarly convey the ideal of holding workers accountable. TaskRabbit has long used the language of "social reputation" to frame its rating and review systems. In the early days of TaskRabbit, Leah Busque (2012) advocated that ratings and reviews could contribute to one's "social reputation" and "inspire more trust." Her idea was somewhat inspired by online reviews on e-commerce sites such as eBay: customers need to believe that an unknown service provider is "trustworthy" before "deciding whether to transact with someone online." When customer reviews are "honest, factual and focus on the performance of the ask" (TaskRabbit, n.d.-c), such information can help other customers to select the best Taskers for their task. Instacart (2020c) suggests that the rating system "helps shoppers continue to provide high-quality service" by allowing customers "to share details on what could have been better." By connecting shoppers' ratings with priority access to batches, the system also "celebrates high-quality shoppers who continuously provide great customer experiences." DoorDash (n.d.-b) also argues that "[r]atings are critical to the success

of the DoorDash platform because they allow [c]ustomer to help ensure that the deliveries completed through the platform are high in quality." DoorDash explicitly states that delivery drivers who typically have low customer ratings "are consistently late, deliver orders with missing items, or behave unprofessionally"; therefore, the platform needs to deactivate these drivers.

Inherent in all of these discourses is that customer ratings can promote high-quality services simply because customers can rate their workers and such ratings or reputation matter for workers. An underlying premise is that customers rate their experience with workers in "honest" and "fair" manner. This is in line with how platforms may strategically frame themselves as "neutral" actors in public discourses (Gillespie, 2010). Customer ratings, together with constructive feedback, should ideally provide trustworthy assessment of service quality and hold workers accountable, especially considering that platform owners cannot monitor all service interactions. From the perspective of platform owners, the deployment of customer ratings seemingly results in a self-regulated environment where workers are motivated to improve their performance and provide safe and reliable service in anonymous marketplaces. Ratings become proxies for trustworthiness and service quality.

One might wonder how platform owners framed other kinds of behavioral metrics such as acceptance rates and cancellation rates. Indeed, platform owners describe the methods of calculation, but do not go further to explain why they are implemented. The general idea of these system-generated behavioral metrics is to facilitate *efficient* service transactions. For example, Lyft (n.d.) suggests that acceptance rate can "help keep passenger wait times short," whereas TaskRabbit

encourages Taskers to accept task invitations to avoid having clients to "start their search all over again" (The Burrow, 2018). Together with customer ratings, or what Uber and Lyft call "two-way" ratings, all of these metrics are framed as boons for the "entire communities" rather than a means to manage workers. TaskRabbit, for instance, suggests that metrics can allow the platform to match and recommend the "right" Taskers to the customers (Wired, 2018). From the perspectives of Uber, Lyft, and DoorDash, metrics can eventually promote a safe and secure environment for both workers and customers. Additionally, metrics can enable workers to identify opportunities for improvement, while allowing customers to enjoy high-quality experiences. As metrics can satisfy the needs of workers and customers, they can ensure the success of the community.

Of course, platform owners can incorporate metrics as part of labor processes with good intentions. Yet, while metrics may ideally promote mutual care for workers and customers, platform owners simultaneously distribute the duties of evaluation and management to customers. Labor platforms ostensibly are not responsible for generating metrics, but invite customers—and occasionally workers (i.e., Uber and Lyft drivers)—to monitor service interactions. This is exemplified by DoorDash's (n.d.-b) deactivation policy, delivery drivers "are responsible for monitoring" their ratings and completion rates. Overall, this section shows how metrics might address the needs of platform owners, customers, and workers.

Summary of the Chapter

In this chapter, I have briefly described the development trajectories of each labor platform and explained the role of metrics in shaping platform-based labor

processes. Specifically, I tried to detail the steps workers must take in order to complete an order on each of the platforms to help contextualize workers' daily practices. I also paid attention to how platform owners designed their platforms in ways that might affect workers' earnings and uses of the platforms.

I have also offered a more detailed comparison between the three cases, by describing (a) the types of platform-based metrics; (b) the methods of calculating customer ratings; (c) dispute process; and (d) the connections between metrics and algorithmic direction and discipline. At the same time, I briefly discussed how these platforms framed their metrics in similar ways. This discussion reveals the promising aspects of metrics on a discursive level. The next two chapters instead focus on how workers interpret and manage metrics and demonstrate how metrics are "imperfect" in practice.

CHAPTER 5

THE POWER OF METRICS AND CUSTOMERS

As I discussed in the previous chapter, labor platforms cast metrics as techniques of promoting accountability and mutual care. The management *by* customers through metrics, particularly customer ratings seemingly create a "win-win-win" situation for platform owners, workers, and customers. This discourse paints a typical image of service management, that is, evaluation enables customers to receive the best service, empowers workers to uphold customer service values, and allows companies to attract customers (Korczynski, 2009b). Management is a conflictual and co-operative process (Hyman, 1987), especially when considering that gig workers lack organizational support. To put this simply, employers always encounter a "dilemma of achieving a balance between managerial control and worker self-direction" (Fuller & Smith, 1991, p. 4). How do, then, gig workers make sense of the metrics they encountered and received in their daily work practices?

In this chapter, I focus on the power of customer ratings, but I also discuss why other forms of metrics about behavioral measures (e.g., acceptance rates) seemed less important to my interviewees. First, I engage with the affective power of metrics (Beer, 2016; Espeland & Sauder, 2016) to describe and explain how metrics could create affective feelings across the three cases. Specifically, participants expressed frustration and anxiety over their metrics, especially when they received unfair or negative customer ratings. As I discussed in Chapter 4, customer ratings are often connected with algorithmic direction and discipline. Participants, therefore, were under pressure to score well, even though they did not see "real" benefits or monetary

compensation for maintaining a good rating. Workers, especially in the cases of ride-hailing and delivery platforms, can lose their ability to continue working on the platform, if their ratings fall below a certain threshold.

Yet, interviewees often lacked sufficient actionable information to make sense of their ratings. They also had limited power to control the factors that might influence their ratings and dispute the unfair ratings. As such, customer ratings may rationalize and reinforce power imbalance between platform owners, workers, and customers. Indeed, not all metrics mattered from the interviewees' perspective. Therefore, I discuss the possible reasons why certain interviewees paid less attention to particular forms of metrics. Lastly, I summarize the findings and reflect on how metrics become a form of algorithmic labor control in the digitally-enabled service workplaces.

The Affective Power of Metrics

Following Beer (2016), I consider metrics as "affective measures" that "are aimed at stimulating anticipation and uncertainty—often coupling these with senses of insecurity and precarity" (p. 210). Metrics are affective because they create social pressure for the subjects of evaluation to compete with their peers and improve their performance (Espeland & Sauder, 2007; 2016; K. Levy, 2015). Compared to rankings (e.g., law school rankings in Espeland & Sauder, 2007), platform-based metrics do not explicitly result in comparison and competition because they are largely invisible to their peers and customers. None of the interviewees discussed how they would directly compete with other gig workers in their everyday life. Nevertheless, metrics still create hierarchies among workers by setting normative standards about "average" and "above average" ratings, which further discipline workers' behavior.

During the interviews, participants expressed affective feelings about the ratings they had received because they considered them as representation of their service quality. To be sure, such affective responses could be positive or negative, depending on the ratings they received. Kuzma who worked on TaskRabbit, for example, told me that he always wanted to read the positive reviews and ratings immediately because he considered as "instant gratification." Jeffery and Maverick both worked on TaskRabbit and considered their status of "Elite Tasker" as a recognition of their services. Similarly, Ronnie, a DoorDash delivery driver, appraised the "Top Dasher" status as "a true recognition of your performance" and "it gives you a satisfaction that you've done well." Here, the status is equated with a "good" worker across these platforms.

Such positive affective responses often came with their assessment of the "usefulness" of metrics. For Jeffery, he appreciated the "Elite" status in part because his profile "will show up for the clients to hire" him. This assessment of TaskRabbit ratings and rewards program was common for the Taskers I interviewed. Murdock, for instance, suggested that "With TaskRabbit, the metrics are very important. I have to have really good reviews … And I've had clients choose me because of the kindly written review that was left by a previous job." While the Elite status did not necessarily lead to higher earnings, Scarlett believed that "people who are experienced with TaskRabbit" would filter "Elite Taskers," so her profile could come up in the search result. Renee was an "Elite Tasker" specializing in cleaning tasks and also told me: "I have a built up profile with reviews and really high metrics. So it becomes … just an opportunity for me to make money and let clients come to me." Renee had 4.9

ratings and 100% of acceptance rates, reliability rates, and response rates. Yet, she still paid attention to her ratings to hold herself to "a high standard" because she believed that when her metrics are higher, she is more likely to get hired on TaskRabbit.

Alongside the increased visibility of their online profiles, the interviewees who worked on TaskRabbit suggested that metrics and rewards program did not have any "real" benefits. It means that they could not help workers to make more money. Kooper, for instance, mentioned that "there's no real incentive getting to be part of this program" because "I don't get more money." Maverick, Kuzma, and Audrey also did not notice any financial benefits from participating in the rewards program. As Kuzma stated, "My rates don't automatically go up. I still have to manually do whatever I want." Even without the "Elite Tasker" status, Taskers could still adjust their rates. Kuzma added that TaskRabbit has a Facebook group for all "Elite Taskers," but the company did not seem to respond to any posts in the group. Even without the financial incentives, the interviewees were still expressed the importance of maintaining their ratings to enhance the visibility of online profiles because they were under the impression that their ratings and visibility of their profiles were essential to clients' hiring decisions.

Ride-hailing drivers and delivery drivers shared similar concerns over the "usefulness" of metrics and rewards programs. Harvey was the only delivery drivers who suggested "when customers see that their driver has a high rating, there's a possibility that you receive a tip." The rest of the delivery drivers, by contrast, stated that metrics and rewards programs did not have any "real" benefits. Carter, for example, stated, "Top DoorDasher is no different than the regular Dasher … At first, I

thought it was like, you will earn more money, but unfortunately you're not earning more money than you should be." Inherent in Carter's comment is that DoorDash should provide monetary rewards for delivery drivers who maintain a high rating to recognize their efforts. Lorenzo found that many of the benefits of Uber Eats Pro program were not available to his area and said, "It doesn't really affect me. We're supposed to get free food and drinks, but it'll pop up and I'll click on it and there'll be nothing."

Likewise, when asked about Uber's rating system and virtual compliments (e.g., badges), Austin shared, "I pay attention to it. You know these compliments on the app, every once in a while, I get a compliment, it just kind of makes me smile. It's love. It's good to have them, but it doesn't really make a difference." Here, Austin had a more critical view of the "usefulness" of such metrics. He added, "The rating system is more or less kind of useless. It doesn't really serve a purpose." This is because even if he maintained a high rating, he would not get higher pay per rides. Noel echoed, "The most frustrating thing is that I can please all my customers, and I go above and beyond, and don't want no pat on the back from Uber, but I'm never recognized for it." The comments from Austin and Noel reveal their critique of not getting recognized for outstanding service which was exemplified by their high customer ratings.

Additionally, other interviewees felt anxious about their ratings, in part because they were well-aware of metrics-related rewards and sanctions. When asked about his experience of Uber's 5-star ratings, Dominic connected ratings with the Uber Pro program. He explained,

The two biggest things that matter are if you're diamond, you get preferred

> pickups at the airport. If you take someone to the airport, you get matched on a ride faster and you get the destination. If we're closer to the end of the ratings year, which actually one just ended January 31st, I would be more anxious about it. Because you need 4.85 to get it. So I'm generally above it, but I just want to make sure that I can at least be, especially with Uber, be able to see every big event for maximizing my earnings. So I want to keep that.

Being a "diamond" driver enabled him to get more information about the estimated direction and time. He considered such information as "the only real tangible benefit" to maximize his earnings. He was under pressure to maintain a 4.85 average rating to keep him qualifying for the rewards program.

Jeremy—an Uber driver working on the platform for four months—discussed the disciplinary outcomes of ratings. He did not frequently check his average customer ratings, unless his ratings decreased. As he described,

> I did look at them [i.e., customer ratings] when I've noticed that my ratings start to go down for some reason. Two months ago, my rating went down from 4.7 to 4.6 and I was freaking out because I knew if my rating kept going down, I could potentially be deactivated.

Jeremy paid attention to his ratings, as Uber sent him a text message to warn him that he had a low rating and, therefore, was at risk of being deactivated. Recalling his feelings of receiving the warning message, he shared, "I was extremely confused, and I was panicking because well if they deactivate me, it kind of makes me screwed." In addition to his customer ratings, he talked about the problems of acceptance rates. He said, "So, in a way, it's kind of like I should be an independent contractor as an Uber

driver but if I reject too many ride requests, that could put my standing on the platform in jeopardy." Inherent in Jeremy's comment is the fear of being punished by the platform due to his metrics.

Indeed, the risk of being deactivated was the primary reason driving the interviewees, particularly those working on ride-hailing and delivery platforms to monitor their metrics. Shawn, for example, only looked at his customer ratings on DoorDash "because if [I] get under a certain rating, then that's when they terminate me." Dylan, too, was concerned with the satisfaction rate on Uber Eats to avoid being "eligible to be kicked off." Like Shawn and Dylan, Sean mostly monitored his completion rate on DoorDash to avoid being deactivated, but he also expressed frustration over the tension between maintaining a good completion rate and maximizing his earnings. He offered an example:

> So, for example, something I do pretty often is I'll accept an order to see how much money the order was because people who order more money worth of food are usually going to tip more. If I accept an order and because it says it has 12 items and I pull open the order and I accept it and it's 12 water bottles, it's not a big order, it's probably not going to be a big tip. So it probably won't be worth the time. I'll cancel that order. That affects my completion rate. Like I said, if the completion rate drops under 80%, you could be deactivated. The one thing I monitor pretty constantly is my completion rate, just to make sure I don't dip under that number when I'm trying to be more strategic.

As DoorDash withheld actionable information about delivery orders from delivery drivers, Sean had to accept an order to look at his actual earnings. Yet, once he

accepted the order and decided not to perform the unprofitable order, he would need to relinquish his completion rates. His experience exemplifies how metrics become intertwined with other parts of labor process; specifically, even though platform owners did not provide sufficient information about the order (i.e., algorithmic direction), he must accept most of the orders he had received to maintain the completion rates.

Unlike ride-hailing drivers and delivery drivers, Taskers would not be deactivated because of their customer ratings, but they still expressed frustrations when receiving a negative rating. Audrey worked on TaskRabbit for one year and mainly performed cleaning tasks. Similar to Jeremy, she looked at ratings when they went down. She stated, "while it's down, when it looks bad, that reflects or impacts your ability to get hired. So that can be super frustrating." Here, ratings mattered because they could affect one's employability on TaskRabbit. Kooper, similarly, talked about the pressure of maintaining his ratings and reliability rates. For him, working on TaskRabbit was similar to running a business. He had to manage his availability and know his "own threshold limit." He stated, "

> There were times to where I might have burned myself out on a day or two. You can't anticipate the depth or the level of work or certain tasks, but there are definitely times where I might have booked myself wall to wall over a weekend … The consequences of do I still do those tasks because the money is needed, or do I suffer? Ratings dip and reliability dip if I cancel it.

Kooper's narrative focused on the work of scheduling tasks on the platform. When he was over-scheduled, he still had to perform all of them not only for the earnings, but

also for maintaining his ratings and reliability rates. It is noteworthy that scheduling tasks is a key component of Taskers' everyday life, which distinguishes Taskers from workers in the other two cases. Taskers can schedule their tasks and communicate with their customers in advance, whereas ride-hailing drivers and delivery drivers can only accept jobs that are assigned by the platform. As I will discuss in more detail in Chapter 6, Taskers must communicate with their clients to manage their expectations about the tasks, which also allows Taskers to manage their customer ratings directly. While Taskers' clients would estimate the length of their tasks, such estimates were often inaccurate. Yet, even they did so, customers could be late or suddenly change the time of the tasks. Therefore, all of the interviewees avoided scheduling back-to-back appointments because if they were late on the previous task, they would need to contact the next customer to reschedule or even cancel the appointment. This could negatively affect their customer ratings and reliability rates.

Taken together, customer ratings became "affective measures" (Beer, 2016; see also Espeland & Sauder, 2016), even though workers did not find any monetary benefits associated with their ratings and participation in the rewards programs. While Beer (2016) offers a conceptual discussion of how metrics exercise power by provoking people's feelings, this project provides empirical evidence to demonstrate how customer ratings could induce anxiety, even without direct competition among workers. Although none of the interviewees had a low average rating that could lead them to lose employment opportunities, the linkage between ratings and discipline generated affective responses. What distinguishes this project from previous studies of law school rankings (Espeland & Sauder, 2016) is that gig workers often lack

resources to navigate ratings and platform-based customer ratings lack clear amenable aspects.[22] Situating into the three-way contest for control, the presence of customers further complicate workers' affective responses. From the perspective of platform owners, ratings could motivate workers to make customers happy, or to provide what they considered "high-quality service." But the presence of customers itself is a source of uncertainty because platform owners cannot require all customers to use the same evaluative standard to rate workers. As I explicate below, even though workers wanted to improve their ratings, they had difficulties in interpreting and managing their ratings, in part because of the lack of feedback from their customers.

Information Asymmetries and Metrics: "I don't really know …"

The visibility of actionable information matters for workers' decision-making capabilities and autonomy (Braverman, 1974; K. Levy, 2015; Moore, 2018). As in Sean's comment about the tensions between maintaining completion rates and maximizing his earnings, we can observe how frustration over metrics could arise from the lack of actionable information before accepting an order. In this section, I focus on the lack of information about metrics for workers.

In an *ideal* scenario of service management, workers could learn from customer feedback and the direction of their managers. Yet, both feedback and direction were largely absent across the three cases. One way of thinking about this issue is whether customers tend to leave positive or negative feedback. Previous

[22] This is not to suggest that law school rankings do not induce uncertainty and anxiety. But law schools could hire administrators to monitor the rankings of themselves and their competitors (Espeland & Sauder, 2016). Meanwhile, law schools could strategically manage their graduate employment data, enrollment data, and faculty leave schedule to increase their rankings.

research on online review systems on eBay and other marketplaces identified a positive bias of feedback as the other party would retaliate against negative comments (Askay, 2015; Dellarocas & Wood, 2008). This may not be useful for explaining the case of labor platforms, in part because workers cannot rate their customers (TaskRabbit, DoorDash, and Instacart) and in part because even if workers can (Uber, Uber Eats, and Lyft), customers would not see an immediate change in their ratings. Nevertheless, when I asked interviewees to describe their metrics, they claimed that many customers did not rate them. Scarlett, for example, commented, "I can go a week where it's like everybody is suddenly giving me a review and then go another week where nobody gives me a review. So, it seems pretty random." Kuzma also told me that he had 59 five-star reviews on TaskRabbit, but he had done hundreds of tasks. Noah, similarly, completed over 2,000 rides, but 1,000 of them had no ratings on them. Daniel, another Uber driver, also stated, "Not everybody rates me. I would say, probably 40% to 50%, maybe even less than that." Comments like this indicate that workers seldom got reviewed.

As non-rated tasks do not count towards workers' overall ratings, this does not seem to be a problem. Nevertheless, even when workers did get reviewed, the information was seldom actionable. What becomes problematic, as exemplified by Renee's observation, is that "a lot of people will just give you a star rating and not take the time to actually leave a review, but it still shows up and still counts." For Taskers, a rating counts towards their total as long as they remain working on the platform. When asked about his experience with negative ratings on TaskRabbit, Kuzma shared,

> Particularly the ones that don't offer an explanation, those are the most frustrating because I'm a person to admit that I don't know everything or I'm not the best of everything. There was something that you didn't like and especially if you didn't communicate it to me verbally in person. The least you could do is to say why you weren't happy with this particular service.

Neyland, too, expressed that his frustration because he had never received any written feedback on his 4-star reviews. Timothy, an Uber and Lyft driver, also stated that he did not really know why he received a 1-star rating because the rider did not provide any explanations for the rating. These comments indicate that the lack of qualitative feedback from customers made it difficult for workers to make sense of their ratings. That said, workers wanted to have customers' explanations for the ratings below 5 stars because platform owners set a high rating threshold for the continuity of employment (e.g., Uber's 4.6/5.0 threshold for deactivation) and platform-based rewards programs.

Another way of thinking about information asymmetries is that workers did not have sufficient actionable information to interpret and manage their metrics. There was a common refrain among the interviewees that they "don't really know." Dominic's observation about Uber's rating system exemplified this epistemic uncertainty:

> You don't really know because a lot of times I won't even know when I got negative ratings until later. The only negative ratings, because sometimes when someone rates you low, sometimes they'll send you emails to say, "This passenger said you were rude," or, "This happened, the fact that you were

> driving aggressively" … I don't even know specifically why I got a negative rating. Also, a lot of times they're just confusion because a lot of the comments will be like, "A rider said that you were driving aggressively. Do you remember this incident?"

Uber drivers cannot see the individual rating for each ride; instead, they can only see changes in their average ratings and riders' compliments for their positive ratings in the driver app. But such changes may be delayed because riders are not required to rate drivers immediately after the ride. Also, when drivers look at the "rider feedback" in the driver app, they may only see the "top reported issue" such as "cleanliness" without any further explanations. Hence, Dominic did not know whether riders left any specific feedback for a particular ride. While Uber sent weekly email to summarize the performance of drivers, Dominic only received a brief explanation about his ratings. This brief explanation, however, was not useful for understanding why he had received a particular negative rating. Here, it remained unclear how Uber processed customers' written feedback and made them visible to workers. Recalling his experience as an Uber driver in 2014 and a Lyft driver in 2017 and 2018, Eddy shared, "When I got low ratings, I very often had no idea why. You'd see your average drop and you'd wonder what happened. I don't like the anxiety of that." Austin, meanwhile, reflected on the negative rating and complaints he received on the Uber app: "I've gotten emails from Uber saying that you've been flagged for discrimination or something, and I don't really know. I'm a very friendly driver, so sometimes I'm just like, 'I don't really know.'" Noel added that there may be time lag between customers' act of rating and the reflection of ratings on drivers' app, because "our

ratings don't always show up till maybe two or three days, or even a week later."

Carter's critical reflection on DoorDash's rating system further suggested that it should be the platform's responsibility to make clearer the signals of metrics, rather than relying on workers themselves to figure them out. As he offered:

> One thing that I'm very concerned about it is I understand that DoorDash gives us the ratings, I understand that they're telling us what is the range between 4.60 is medium, or 4.80 is high. 5.0, it's higher, above average. But what I really would like for them to do one thing, and this is what I requested numerous times. Please, if our customer ratings drop, they should be giving us a list to know, what was going on, why did our rating drop? For example, I'd like to know, if maybe one of the orders was not complete, so I know how I can do [a] better job next time. And that's a big problem because for us to figure out what's really going on, DoorDash is not giving us that hint.

Inherent in Carter's comment is the difficulties in following the evaluative standard set by the platform. Throughout the interview, Carter seemed to agree that customer ratings could maintain service quality and suggested that there should be "ratings for customers" to hold them accountable. While DoorDash asks customers to rate the worker and restaurant separately, as I illustrate below, the interviewees were under the impression that customers did not do so in practice. For Carter, DoorDash should take a step further to give workers detailed feedback to improve their performance.

So, what should such "detailed feedback" look like? During their expositions, participants mentioned that they would like to know more about *who* gave them a negative rating (i.e., the cases of ride-hailing and delivery platforms) and *why* they did

so (i.e., the three cases). For reasons of privacy, it might not be feasible for platform owners to disclose the identities of customers. But from workers' perspective, they wanted to know the identities of customers in order to recall specific service interactions. This would also help workers to better understand the reasons behind negative ratings and even to consider disputing unfair ratings. That said, workers did not necessarily oppose all forms of evaluation. As in Carter's earlier comment about DoorDash's evaluative standard, he would agree that ratings could promote accountability, but it was difficult to realize the full potential of ratings without actionable feedback from the platform and customers. It is against this backdrop that I argue metrics, particularly ratings, are "imperfect" in practice. The epistemic opacity of metrics could be *unintentional* when considering that customers simply do not leave any qualitative feedback. Workers' autonomy is constrained by the difficulties in understanding their ratings.

Out of Workers Control: "There's nothing I can do about it"

Even though the interviewees did not have much feedback about the negative ratings they received, they would speculate possible reasons based on their work experiences and the hints from the platform. Here, the speculation was more about their service interactions with customers and the ways that customers used metrics to evaluate their performance, rather than about the inner workings of algorithmic metrics. As in the stories of Carter, Jeremy, and Kooper, they were keenly aware of disciplinary outcomes of metrics. Carter's comment made it clear that he did follow DoorDash's evaluative standard to think about average and above average ratings. This is not to suggest that interviewees fully understood the intended design of

metrics, but to highlight how they attempted to find ways to interpret their negative ratings to avoid sanctions. Nonetheless, many of the factors my interviewees—and online discussions from worker-to-worker communities—identified were out of workers' control, or as Lorenzo put it, "there's nothing I can do about it." What is at stake here, as I illustrate below, is the limited worker power to modify their performance to play by the anticipatory rules of ratings.

Of course, anticipatory rules of ratings evolve as platform owners change their policies and technological design of the infrastructures. As I began studying ride-hailing platforms in early 2017, Uber and Lyft had not implemented what they called "ratings protection"; that is, platforms would automatically remove negative ratings if riders selected a reason that is out of drivers' control. This kind of ratings protection, as I discussed in Chapter 4, becomes commonly adopted by ride-hailing and delivery platforms. But back in early 2017, discussions from ride-hailing drivers' online communities—such as UberPeople[23]—often referenced bad weather, heavy traffic, and navigation. My interviewee Noah echoed this and considered them as "false ratings" because, for example, drivers could not make sure that Uber's in-app GPS system would provide the most accurate location and route. Howell, too, suggested that some of his negative ratings had to do with Uber's navigation system, but Uber eliminated such ratings as the company implemented the "180 Days of Change" campaign.

The interviewees working for ride-hailing platforms suggested that the

[23] Although the online forum is called "UberPeople," participants can discuss all of the ride-hailing platforms, including both Uber and Lyft, on the forum. This is also one of the largest online forums for ride-hailing drivers.

platforms, particularly Uber should educate riders about their rating. As Uber drivers are at risk of being deactivated if their ratings are below 4.6. One to four stars are considered failing grade. Cohen, for example, considered that ratings could hold "drivers and even riders accountable," but he proposed that "there's a component of education that passengers need to understand what's a good rating, what's [the rating] drivers can potentially be deactivated for." Noah further denounced the unfairness of Uber's rating system. As he explicated:

> The average person, they think that three stars is average, four stars is good, five stars is excellent. So if you just give a basic ride to somebody, you don't do anything over and beyond, you pick them up, you take them somewhere, you drop them off, the ride is over, you're given three stars, it was an average ride. The problem is if you receive 100 four-star rides, you would be kicked off the platform because it's not 4.6. You would have to consistently deliver five-stars every single time.

Noah's comment, again, indicates how disciplinary outcomes direct drivers to continuously monitor their performance and consistently get 5-star ratings. His criticism of Uber also reorients the responsibility for oversight from customers to the platform. While riders are tasked with rating drivers, they are not instructed on the evaluative standard for doing so. Additionally, it is noteworthy that my interviewees were concerned less with Lyft's rating system, in part because Lyft would automatically rate drivers 5-stars if riders do not leave any ratings and remove the single lowest rating for every 100 rides. Considering that many of the riders do not rate drivers, it becomes easier for Lyft drivers to maintain their ratings.

Perhaps not surprisingly, as customer ratings are essentially about evaluation of service interactions, many of the interviewees speculated why customers might be dissatisfied with their service, even though the situation was out of their control. In the case of delivery drivers, interviewees stated that customers might give them a low rating primarily because customers had a long wait for the order or there were missing items. While these reasons seemed fair at first glance, it is noteworthy that not all restaurants prepare for delivery order in a timely manner. As Melvin shared, "in terms of being on time or early, that's somewhat out of my control." He explained that drive-throughs at Burger King and McDonald's usually took longer than other restaurants, especially when their lobbies were closed. He tended to go to restaurants where their lobby were open to "speed up the process" but it depended on the order and place. Sean, too, said, "There are restaurants that I definitely avoid because they aren't quite honest with me about when the food will be ready, and I end up waiting there for a long time." Dylan suggested that "Uber Eats drivers may receive lower priority at certain restaurants and they will focus on other customers even if they arrive afterwards." As a result, delivery drivers may have a long wait time at a restaurant. The waiting time, meanwhile, is not compensated.

The interviewees, meanwhile, attributed the responsibility of missing items to restaurant staff, especially during COVID-19. Before COVID-19, DoorDash and Uber Eats suggested delivery drivers to check if the items inside the bag were accurate. During the pandemic, many restaurants seal their delivery bags for safety reasons. Hence, it would be inappropriate for delivery drivers to open the sealed bags to check the order. My interviewees like Ronnie, Clara, and Nelson encountered this situation

and received a negative rating accordingly. As Ronnie described,

> Almost all of the orders are bagged, and they're sealed, and we cannot open those orders. And so, sometimes the restaurants fail to include an item that the customer had requested, or fail to complete a specific request that the customer requested. Like, they didn't want any onions on their burger and their burger has onions. So, the costumer will in turn, when they're asked by DoorDash how they felt about their delivery experience, they will rate us negatively based on what happened at the restaurant, which was beyond our control. That can be discouraging at times, and that can have a negative impact, because it's not a reflection of the DoorDashers ability to deliver the order in a timing manner.

Comments like this show that customers might misattribute the responsibility to delivery drivers through their ratings. Ronnie's experience is also telling of the distinction between evaluation of *delivery experience* and of *restaurants*. Indeed, DoorDash and Uber Eats have two separate rating systems for delivery drivers and restaurants, but they seem to be ambiguous in practice. When asked about his negative ratings on Uber Eats, Charles stated that he did explain to a customer who was angry about the wrong order that his rating was different from the restaurant's one. But immediately after he had completed the order, his rating dropped. Then, he contacted Uber to dispute this unfair rating, but Uber refused to remove it. In line with the previous discussion about ride-hailing platforms' ratings protection, delivery platforms claim that these unfair ratings would be automatically removed, but it is still

dependent on whether and how customers select their rating reasons.[24]

Like the other two cases, Taskers do not have full control over their work schedule, which might threaten their ratings. As I discussed previously, my interviewees tended to avoid scheduling back-to-back appointments because a client could be late or request for changes in the task at the last minute. In addition, (mis)communication between Taskers and their clients can significantly affect Taskers' ratings. As Kooper put it, "it was just impossible to make every client happy," especially when his clients did not communicate their expectations clearly. He offered an example of working "remotely":

> They might leave the keys under a mat or with a concierge. You never see the client in person. So, you do the job and then you leave and then they come and write a bad review, but then you're not able to communicate what was going on. They didn't communicate to you how you could assist. You couldn't defend yourself ... For example, like maybe there's a scratch or a mark that I just cannot get out, I'll go say that, that, "Hey, I tried everything. I cannot get this dirt out for whatever reason." So they are aware of the problems as I go forward. That's as far as negative reviews.

The narrative signals the importance of communication between Taskers and their clients because each client can have very different expectations. Even when a client agrees on the task details, they can be dissatisfied. Zoey, for example, shared:

> They [i.e., TaskRabbit] sent me last week to go to a store and buy someone a

[24] DoorDash, more recently, makes visible how many "unfair ratings" has been excluded from their delivery drivers' profile on the worker-facing app. But this change is implemented after I conducted interviews with DoorDash delivery drivers.

> plunger that was 15 miles away from me. And I'm like, "What? You have a one-hour job to buy a plunger 15 miles away?" So I told the lady I couldn't do it unless you pay me for hours that thing. That's like 50 bucks, and she said yeah, she was desperate. But she didn't tip me, and she didn't give me a good ranking, she didn't give me anything because she didn't want to pay $50 for someone to go buy her a plunger.

For Zoey, distance mattered because clients could hire her to do small tasks (e.g., purchasing a plunger) for one hour. When she accounted for the travel time, such tasks were not profitable. In the above example, she explicitly told her client that she would need to charge $50 and make it a two-hour task. Her client accepted the request, but eventually gave her a negative review. In short, the two examples from Kooper and Zoey affirm that Taskers cannot control how their clients rate them because the platform delegates managerial authority to customers.

According to TaskRabbit, clients can only rate their Taskers *after* the tasks are completed and invoiced. Yet, Easton and Neyland told me that they received ratings from the tasks they had declined. Neyland, for example, said "On TaskRabbit, even if you don't complete the task, they can still rate you. TaskRabbit's perspective or reason behind this is that it also matters on how you communicate." He recalled that one of his 4-star ratings was given by the client "who wanted me to have a pickup truck," but he declined this task because he did not have the truck. Easton suspected that this may be a technical glitch in TaskRabbit's app. On the subreddit for Taskers, I also found a few discussions about this issue, which were generally in line with Easton's suspicion. The members of the subreddit also suggested that Taskers might be able to dispute

these negative reviews, if they reported them to TaskRabbit.

Overall, the findings from the three cases suggest that there are factors that may affect workers' ratings, though they are beyond workers' control. This may speak to a wider concern over how customer ratings place workers in a vulnerable position where they must rationalize such unfair ratings and customer-oriented rules.

Power Imbalance Between Platform Owners, Workers, and Customers

Using the term "customer-oriented rules," I highlight how customer ratings reinforce and rationalize power imbalance between platform owners, workers, and customers in two ways: (a) the experiences of dealing with "difficult" customers and (b) workers' limited power to dispute unfair ratings.

The experiences of dealing with "difficult" customers emerged as a prominent theme about the interviewees' frustrations over their platform-based work. The "difficult" customers generally refer to those who have "unreasonable" expectations about gig workers and do not respect the jobs of gig workers. In the case of ride-hailing platforms, examples of unreasonable customer expectations include requesting drivers to pick them up in places where drivers cannot park and stop, squeezing too many people in the backseat of the vehicle, and asking drivers to go over the speed limit. These requests, meanwhile, can place workers at risk of legal consequences (Maffie, 2020b). Lyndsey's comment exemplifies how ratings could enable customers to punish ride-hailing drivers. She provided an example:

> The lady was late for work. I was there on time and she wanted me to run red lights, make right turns on red when you're not supposed to, [and] wanted me to speed up. You know, breaking laws to get her to work on time. In other

> words, she wanted me to make up [for] lost time. I couldn't do that, so she gave me a low rating.

When encountering such unreasonable requests, my interviewees told me that they would refuse to do so. Lincoln, for instance, stated that he would refuse having riders to squeeze four people in the backseat to avoiding getting into trouble. The riders, however, might still insist on their request for a while, and eventually treated him impolitely (e.g., slamming the car door).

To be sure, not all unreasonable requests are beyond the legal limit. Most of my interviewees talked about how they had attempted to follow all of the customer requests, but still did not receive a 5-star rating. Noah, for instance, told me that he had a rider who gave him a wrong pickup location, which led him to spend a while to communicate with the rider to get the accurate address. During the ride, the rider continuously gave him commands of navigation. He shared, "I felt more like a robot and like somebody being ordered around that someone providing some transportation on where you had to go." He followed all the requests, but eventually received a 3-star rating. He added, "That whole scenario was completely unfair, in my opinion, because I was given faulty information and I was put in a situation where I was highly dependent upon him to give me correct information to provide him the service."

Meanwhile, Collin, an Uber Eats delivery driver, suggested that "the worst customers I get are the customers who don't understand what the job of the driver is." For instance, customers may ask him to "make sure they [i.e., the restaurant] did this to my sandwich," but he could not open the sealed bags to do so. As a result, customers might give him a low rating because of the restaurant's oversight. Or, as

Zoey explained, “the worst thing is the people that rate you because you always have to do whatever they say.” Specifically, she was hired for an hour to pick up balloons and a cake for a birthday party, but she eventually had to set up the party and clean up afterward. She was paid $20 and received a 1-star review on TaskRabbit. Reflecting on this experience, she continued, “people can misuse you, and you can’t get out of it ... because they have the right to rank you.” Inherent in their reflections is the power of customers through their rating practices.

A potential counterargument, particularly from Uber and Lyft, is that drivers *can* and *should* use their “two-way” rating systems to hold these customers accountable. From the interviewees’ perspective, the so-called “two-way” systems did not seem to address the issue, but rather to rationalize power imbalance between workers and customers. Platforms seem to rarely punish customers, as illustrated by Noah’s observation: “A customer can have a rating as low as possible, and still get a ride. I mean, I’ve picked up somebody who had a 3.2 rating … You’ll never see a driver with a 3.2 rating.” Tellingly, he added, “Why is it customers can basically do whatever they want to do, but the drivers can’t? Well, the reason why is simple: The customer is the person who has the revenue … not the drivers.” While Uber’s business model remains an interesting empirical question, his critical observation highlights how the company tends to favor customers over workers. Reflecting on his ratings on restaurants and customers on Uber Eats, Collin noted, “the thumbs up and down for the drivers has a very real effect, but I don’t think it has much of an effect on the restaurants and the customers. I don’t even think the customers know they get rated.” Despite the unequal disciplinary outcomes for workers and customers, a small subset

of the interviewees would use riders' ratings as a reference point before service encounters, which I will discuss in Chapter 6.

One might also reasonably wonder why gig workers do not dispute unfair ratings to labor platforms. Except for platforms' weekly emails, workers have limited communication with platforms. Recalling his experience of contacting Uber and Lyft, Adrian shared that the help desk simply read "you a basic script they have followed, which they don't understand day-to-day" operation. Collin also raised this issue even though he had Uber Eats' priority support number, due to his rewards program status. He stated, "Uber Eats support, I think, is meant more the customer than for the driver because when I talk to the support, I describe my problem to them and they have no idea what I'm talking about."

The interviewees' attitude towards platform-based dispute process, by contrast, was quite positive on TaskRabbit. They were generally satisfied with the *outcomes* of the platform's support, which I will explicate in the next section. But, as Audrey and Kooper pointed out, the process itself took a long time because they had to use email to communicate with the company. There is no phone number for Taskers to report problems in real-time. Kooper explained that TaskRabbit usually would respond to his email within 48 hours, but it might take several emails to tackle the problems. Therefore, he learned to "become very detail-oriented" to document the issues to prevent any confusion in this process. All these experiences are indicative of the additional, unpaid time and efforts workers must spend if they want to dispute an unfair rating and report problems to platform owners.

Importantly, platform owners seemed to always take the side of the customers.

This response emerged from the interviewees' reflection on how platforms handled their complaints. For instance, Clara once received a negative rating from the customers because of a missing delivery order, even though she remembered that she had delivered the order. She checked the app and found that DoorDash reimbursed the customer. She told me that DoorDash might check with the customer and restaurant to investigate if the order was missing, but this process did not involve her. As she shared:

> You are in the middle because there's nobody who will be stand behind you. You have no testimony. You have no witness. You have nothing. It's your word against two of them [i.e., restaurants and the customer]. It's not too many merchants who they're [i.e., restaurants] going to recognize when they make [a] mistake.

Inherent in Clara's reflection is a sense of powerless for workers to dispute unfair ratings. Or, as Collin put this, "I can afford to fight a restaurant worker because Uber sees both of us as sorts of equals, but if I fight a customer, forget it, even if I'm right, Uber will deactivate me."

Derek, meanwhile, condemned Instacart for enabling "customers do anything" to "hurt the shopper." He offered an example:

> We do take pictures when we deliver. And many times, we take pictures of delivery or the good, and then if a customer claim they have a wrong item, totally fraud, and we can report it in the app, but that doesn't remove the rating. So Instacart will say, "Well, we will look into it," but the rating is not removed for the time being. So, let's say Instacart spends five days to

> investigate, but that 1-star rating will remain on your record and people will need to do another hundred orders to remove that fake 1-star rating.

As Instacart shoppers' ratings determine their access to batches, it would take a long time for Derek to recover from the 1-star rating. Tellingly, he added, "the customer has total privilege to give you any ratings they like."

What is at stake here is that customer ratings—and, relatedly, the ways that labor platforms manage their workers and customers—may rationalize and reinforce the logic of following customer-oriented rules. Customer ratings, therefore, are imperfect systems that incite workers to accept any ratings from their customers, even though such unfair ratings could threaten their employment opportunities.

The Varying Disciplinary Power of Metrics

Much of the previous discussion focuses on the convergent themes about the power of customer ratings. It would be too simplistic to assume that all forms of metrics affect gig workers equally. Two questions, therefore, remain unanswered: How about the power of other forms of platform-based metrics? Are there any differences across the three cases? To address these questions, it is vital to pay attention to those who were less responsive to metrics.

The interviewees paid less attention to system-generated behavioral metrics for two possible reasons. First, these metrics generally have limited disciplinary outcomes Derek, for example, explained that cancellation rates and speed metrics on Instacart "are not important" because they do not "impact how many batches you will see." Nelson, meanwhile, only looked at his ratings to avoid being deactivated by DoorDash, but he did not care much about his acceptance rate, competition rate, and

on-time/early rate. He told me that his acceptance rate was about 30%, meaning that he would not be qualified for the "Top Dasher" program. But he also did not think "the benefits of being a Top Dasher outweigh of just being critical with what orders you accept." For him and all of the interviewees, making money was the primary reason driving them to work on these labor platforms. The second possible reason is that these metrics were easy to maintain, particularly in the case of TaskRabbit. In this case, the majority of the interviewees (7 of 11) had 100% of acceptance rates, reliability rates, and response rates. Such metrics do not involve human assessment and seem to be easy to maintain. This is also because Taskers can specify their skills and communicate with customers before accepting a task.

To be sure, a small subset of my interviewees told me that they had cared less about their customer ratings over time, or never looked at their metrics. Brandon and Dominic explicitly articulated why they still looked at their customer ratings but tended to care less about them. Brandon tended to care about negative customer ratings in his early days of work, but he later found that a few negative ratings would not hurt his status as a "diamond" Uber driver (i.e., his current rating was 4.9). Dominic's reason, by contrast, was related to his economic dependency on the platform. He explained, "Now I care less about my ratings. Now that my livelihood doesn't depend on it [i.e., Uber] … I don't look at it [i.e., ratings] as much." When he first started driving on ride-hailing platforms, he was in the graduate school and this platform-based work became his primary source of income. As he graduated in 2019, he decided to continue working on the platform as a flexible, part-time job.

Among all of the interviewees in the three cases (n = 50), Joseph and Sofia

never looked at their metrics, whereas Shawn, Roger, and Harvey monitored their customer ratings but were not familiar with the rewards programs on their platforms. Except Roger who worked on Uber Eats and DoorDash for one year, the rest of them began working on as delivery drivers after August 2020. At the time I interviewed them, Roger, Harvey, Shawn, and Joseph were students, whereas Sofia had a full-time professional job. None of them considered working on the delivery platforms as their long-term job. They (except Roger) were arguably newcomers in the gig economy and did not plan to stay in the future. This may explain why they were less familiar with platform-based metrics and rewards programs. Joseph, for example, did not know there were metrics about his performance on Instacart. This is consistent with existing research in which workers' economic dependence on the platform could inversely influence their degree of autonomy (Schor, 2020).

Lastly, my interviewees who worked on TaskRabbit seemed to be more satisfied with the platform than those who worked for the other two types of platforms. This is similar to Schor's (2020) finding that TaskRabbit workers had a high satisfaction level. Five of the interviewees explicitly considered TaskRabbit as a "good company," primarily because the company had relatively good communication with them. As previously discussed, email exchange between TaskRabbit and workers may take a while. Nevertheless, Jeffery also said that TaskRabbit would help him to remove an unfair rating. As Zoey described, "they always fix," as long as she sent all the documentation to TaskRabbit. This is also in line with Kooper's narrative in which he learned to be "very detail-oriented" during his communication with the company. By contrast, the other two types of platforms generally had stricter dispute policies.

Uber clearly claims that no individual ratings can be removed upon workers' request, whereas the other platforms seem to allow workers to dispute an unfair rating based on their policies. None of my interviewees, except for those who worked on TaskRabbit, said that the platforms removed the unfair ratings they had disputed. Yet, TaskRabbit is a relatively small platform in terms of the workforce[25] and customers, compared to other platforms. In this vein, the size of the workforce and customers may explain why it would be easier for TaskRabbit to process and handle complaints from workers. Despite the distinction between TaskRabbit and the other two cases, all of these platforms retain the right to decide how to handle workers' complaints and do not disclose much about the dispute process. Moreover, as I demonstrated previously, Taskers still expressed frustration over their customer ratings.

Summary of the Chapter: Metrics as a Form of Algorithmic Labor Control

In this chapter, I have discussed how metrics, particularly customer ratings become a form of algorithmic labor control through the production of uncertainty and anxiety in the three cases. The deployment of customer ratings led gig workers to feel anxious and frustrated, which might further incite them to continuously monitor and maintain their ratings. In this vein, customer ratings facilitated platform surveillance over workers. The affective power of customer ratings became intertwined with algorithmic discipline. Although labor platforms offered rewards programs to workers, the interviewees seemed to focus more on sanctions rather than such virtual rewards (e.g., ratings, badges, and elite status). Workers' employability—including the

[25] This may also be reflected in the number of members of the subreddit for TaskRabbit (r/TaskRabbit; i.e., 2,200), whereas the subreddits for the other platforms I studied had much more members (e.g., 19,200 members in r/lyfdrivers and 135,000 members in r/uberdrivers).

opportunities of getting hired (TaskRabbit) and the risk of being deactivated (ride-hailing and delivery platforms)—produced the affective power of customer ratings. Inherent in this form of platform surveillance is that a high customer rating and worker's rewards status signal their "good" and "professional" service on the labor platforms. That said, metrics and their associated rewards and sanctions set normative standards for service interactions and create hierarchies among workers, even without direct competition among workers.

Moreover, I have explicated how the power of metrics was exercised through the creation of information asymmetries. The interviewees were aware of the disciplinary outcomes of customer ratings, but they did not have sufficient qualitative customer feedback to interpret any low ratings they received. The lack of feedback made it difficult for workers to improve their performance and even dispute unfair ratings. The lack of actionable information also led workers to speculate the factors that might result in a low rating. Many of the factors were, however, beyond their control. Customer ratings not only delegated the managerial authority to customers but also placed workers in a vulnerable position to rationalize customer-oriented rules. The interviewees were conscious of platform owners' unequal treatment of themselves and customers, especially when considering their service interactions and limited communication with platform owners.

In the last part of this chapter, I have identified the factors that might contribute to divergent disciplinary outcomes of metrics across the three cases. First, system-generated metrics seemed to be less important for workers because they were relatively easy to maintain and less connected with algorithmic discipline. Second,

when workers were less dependent on labor platforms as their primary source of income, they would be more likely to ignore customer ratings and rewards programs. Third, the case of TaskRabbit is different from the other two cases because Taskers were quite satisfied with dispute outcomes.

Taken together, customer ratings as a form of algorithmic labor control may exemplify the concept of customer-oriented bureaucracy (Korczynski, 2004, 2007, 2009b). On the one hand, customer ratings enable platform owners to monitor and discipline their distributed workforces at a low cost. On the other hand, workers are under pressure to score well and normalize the punishment of customer ratings. In other words, customer ratings entail both technical and normative control. Situating in the service triangle, platform owners seem to be much more invisible than managers in the traditional service workplaces. The invisibility of platform owners indeed reinforces unequal power-relations between workers and customers. From a practical standpoint, platform owners cannot monitor all service interactions. Assuming that platform owners want to investigate an unfair rating, they must rely on the input from different actors (e.g., customers, restaurants, and workers). The narratives of workers, especially those in the cases of ride-hailing platforms and delivery platforms, indicate how platforms might favor customers (and restaurants) over workers. Despite the difference in the types of service work and the visibility of metrics, workers in the three cases were subject to a certain degree of algorithmic labor control through customer ratings. This contributes to understanding the affective power of metrics in action (Beer, 2016)—even without direct competition among workers, customer ratings can still make workers feel and undermine workers' sense of autonomy in

digitally-enabled service workplaces.

CHAPTER 6

MOMENTS OF RESISTANCE: MANAGING THE "IMPERFECT" METRICS

This chapter examines how workers manage the "imperfect" metrics in their daily work practices. On the one hand, metrics are "imperfect" because they do not provide sufficient actionable information for workers to interpret the meanings, as I illustrated in the previous chapter. On the other hand, metrics can also be viewed as "imperfect" systems for labor control because platform owners cannot monitor all service interactions. This provides space for workers to contest and resist the management *by* metrics and customers through the management *of* customers during platform-based labor processes and information-sharing practices outside of the platform. The moments of resistance may seem subtle and sometimes comply with the calculation logic of labor control, in part because of the risk of being punished by platforms. From the perspective of platform governance, reactivity can certainly be anticipated and scripted (Gerlitz & Lury, 2014). This is in line with the underlying mechanism of affective metrics (Beer, 2016) in which uncertainty and insecurity pressure the evaluated to self-monitor themselves and proactively optimize their ratings (Espeland & Sauder, 2016). Nevertheless, such reactive practices can also enable workers to exercise their agency over platform-based labor processes. Put another way, workers strategically manage their service interactions and the platform to advance their interests.

The moments of resistance involve the management of customers before, during, and after service interactions. Specifically, I will discuss how workers may filter customers *before* service interactions and manage customers' expectations

during and *after* service interactions. Workers carry out the set of practices to routinize their communication with customers and reduce work-related uncertainty around metrics, while performing platform-based work. I also discuss the limits and constraints of these moments of everyday resistance. Then, I document the emergence of worker-to-worker communities. While I largely draw on the findings from the case of ride-hailing platforms, particularly Uber, I also examine how and why the interviewees might take part in online communities for sharing work-related information *outside* of labor platforms. Finally, I discuss how these moments of everyday resistance may disrupt algorithmic labor control.

Management of Customers in the Gig Economy

There are different ways of approaching the relationship between management of customers and workers' autonomy in interactive service work. From the perspective of labor control, managers attempt to routinize service interactions by scripting workers' performance and regulating their emotional display (Leidner, 1993; Hochschild, 1983/2012). In this vein, the routinization of service interactions often seems to diminish workers' autonomy and job skills. The performance of emotional labor may also lead workers to suppress their inner self (Hochschild, 1983/2012). Resistance may be referred to subversive practices that refuse to follow the script; for example, flight attendants refuse to smile, a form of emotional labor required by their companies, to resist organizational demands. Nevertheless, service workers do not necessarily feel alienated and may instead accept their service routines (Leidner, 1993; Korczynski, 2009a, 2013; Lopez, 2010). Workers can find scripting and emotion work useful for them to manage their customers (Leidner, 1993) and even to get customers

to tip (Davis, 1959; Thompson, 2015). A notable example is taxi drivers who have long strategically profiled their customers to maximize their earnings (Davis, 1959). In this vein, workers opt to manage their customers for practical benefits, even though they may still be subject to strict organizational control (Leidner, 1993).

Scripting is much more invisible in the gig economy than traditional service workplaces, largely because platform owners rarely provide training to workers. After the interviewees signed up and got approved to work on the platforms, the only training they had was to watch a series of short videos. Austin, for example, described that Uber asked him to watch an eight-minute video on how to use the app. On TaskRabbit, Kuzma, similarly, shared that "the video sessions are really focused on teaching you how to use the app … because they expect you to know how to do all the skills that you say that you can do when you start the job." That said, metrics, particularly customer ratings become the primary scripting device in mediating workers' expectations of service encounters through their modes of evaluation and discipline. Central to workers' reactive practices, therefore, is workers' understanding of what is being evaluated as relevant and what relevant aspects are more amenable in their situation.

While platform owners use customer ratings to script workers' performance, this chapter directs attention to how workers strategically follow and adapt such scripts to manage customers in order to resist algorithmic labor control. I explicate the management of customers by temporal order and discuss how such practices enabled the interviewees to gain a sense of autonomy over platform-based labor processes. Put simply, the management of customers is about filtering and avoiding what the

interviewees considered "difficult' customers before service interactions as well as managing customers' expectations and needs during and after service interactions. Crucially, my goal is not to argue that all interviewees engaged with such practices, but rather to highlight the potential for these practices to resist labor control.

Managing Customers Before Service Interactions

As I discussed in Chapter 5, dealing with "difficult" customers was a common frustration among the interviewees. The "difficult" customers had unreasonable expectations on workers and tended to rate them low. Central to workers' agency, therefore, is the extent to which they could filter these customers in advance and decide *not* to accept their order. Labor platforms, however, discourage workers to decline too many orders. In Uber's community guidelines, for example, the company suggests that drivers should not discriminate customers based on their demographic characteristics as well as destination and delivery location. Their rewards programs, meanwhile, often have a minimum threshold for acceptance rates.

Ride-hailing Platforms. For ride-hailing drivers, a common strategy was to refuse taking any UberPool rides. The interviewees suggested that UberPool—shared rides—remained a key reason for receiving a low rating, though Uber's rating protection had included "co-rider" as one of the qualifying options for removing negative ratings. Clayton, for example, denounced UberPool as "a bad product":

> I hate UberPool because you consistently get rated lower on UberPool than you do on UberX, because it's a bad product, not because you're a bad driver. No one likes to share a car with somebody. Everyone picks UberPool because it's cheaper, but it takes longer to get to their destination, so they miss their

> movie, they miss work, they're late to work ... If you miss your airplane because you took UberPool, you're going to one-star your driver, even though it's not the driver's fault … Or you'll get in the car with somebody you don't like … Because you're mad about Billy, who's arguing with you about politics or whatever, you had a bad experience on your drive, you're going to rate your driver lower, even though it's not his fault. It's Billy's fault.

Clayton's comment reveals why ride-hailing drivers strategically avoided providing carpooling services, as evident in existing studies (Ferrari & Graham, 2021; Reid-Musson et al., 2020). As there are more riders during UberPool rides, it becomes more difficult for drivers to manage all riders. Accordingly, Clayton decided to "never take an UberPool unless … the surge is so high [that] I don't a choice." Here, the surge refers to Uber's pricing model, meaning that drivers could get a higher rate, when Uber signals that demand for rides increases in real-time. However, Clayton acknowledged that "the surge would only last for 10 or 15 minutes"; therefore, experienced drivers like him would predict surge areas and wait until the surge starts, rather than chase the surge. The technique of predicting surge areas has to do with one's knowledge about the city (e.g., when and where big events start).

Like Clayton, Alexis had negative experience with UberPool. She believed that riders might purposefully not choose UberPool (or "co-rider") as a reason for giving her a low rating. As she explained,

> A lot of the passengers don't understand UberPool … Uber says they put it in their apps to try to protect us from this, but it still happens. UberPool passengers will get in my car, and then they will be mad because they have to

> share the car with other people … They'll be like, "Can you bring me to this place? But we need you to bring us to another place." I can't bring them to another place because it's UberPool. I'm already bringing all these other people to their other destinations. Then they will get in app and they will rate me poorly … They don't say it was an UberPool. They won't say it's because it's an UberPool. They'll just say, "We didn't like her," for whatever reason.

Inherent in Alexis and Clayton's observations is the power of customers in the process of evaluating drivers' performance. Drivers, meanwhile, had limited ability to manage service interactions during shared rides. Perhaps more importantly, drivers had done more work but made less money from shared rides because they had to pick up and drop off multiple riders at different locations. As Brandon put this, "it's a lose-lose for the driver." In order to avoid this problematic situation, the interviewees decided *not* to provide any carpooling services. This strategy, meanwhile, does not violate ride-hailing platforms' guidelines because drivers can select the type of ride-hailing services they want to provide.

Another strategy has to do with how drivers decided whether to accept a ride request based on riders' ratings. This strategy can be considered "gaming" (Sauder & Espeland, 2009), or "manipulation" (Ferrari & Graham, 2021) because Uber and Lyft do not encourage drivers to select their riders based on ratings. Terms like "gaming" and "manipulation" may entail the ethical work of drawing moral boundaries between legitimate and illegitimate practices (Petre et al., 2019; Ziewitz, 2019). I use the terms here simply to suggest how this strategy may break the rules of platforms, but it also simultaneously enables drivers to gain control over labor processes. This strategy was

more commonly discussed on driver-to-driver forums and YouTube channels, but less obvious during my interviews with ride-hailing drivers. On UberPeople, one of the largest online forums for ride-hailing drivers, forum participants contended that drivers should not accept riders with a rating below 4.7/5.0 because these riders are usually "trouble-some" and less likely to tip the drivers.

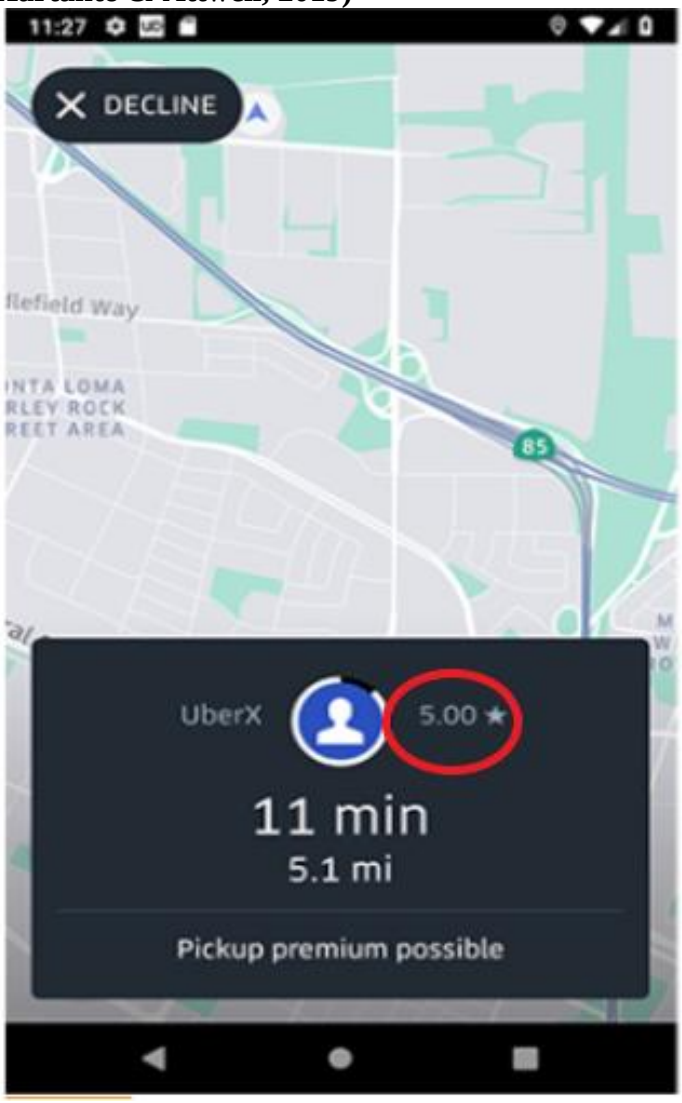
Harry Campbell—who has operated one of the largest driver-to-driver blogs "The Rideshare Guy" and published a book entitled *The Rideshare Guide* (Campbell, 2018)—also shared his own evaluative standard about riders (Campbell, 2015). For him, 5-star passengers mean "someone who comes out within 1-2 minutes of my arrival, has the destination entered into the GPS and is courteous," whereas examples of 1-star passengers are "the puker," "the racist," and "the troublemaker" (Campbell, 2015). He would also give riders 4-star ratings if they made him wait, did not enter a destination, or slam the doors. Ultimately, as he put this, drivers should pay attention to riders' ratings "to take advantage of what little information we get" and "it's advisable to only pick up 4.7 rated and up passengers." Similarly, Bowman (2019), an Uber and Lyft driver in South Florida, also wrote in the *Business Insider* that riders with a rating below 4.7 were likely to be "bad" passengers and he would never accept a ride request from anyone with a rating below 4.6. As he put this, "The lower the number, the exponentially worse a passenger is – rude, louder, dirtier."

I will explain how my interviewees rated their Uber and Lyft riders in the later sub-section about the management of customers after service interactions. When asked about their uses of riders' ratings, only Dominic, Rodrigo, and Sadie said that they would explicitly use riders' ratings as a guideline before service interactions. As

Dominic shared, “On weekends I don’t take low-rated customers because it’s busy enough on Saturday that you don’t really have to. I’ve learned those tend to be the drunker people, the people have the 4.6, or whatever.” Rodrigo and Sadie considered riders’ ratings as proxies for their behaviors but did not use them to determine whether they took a ride. Rodrigo, for instance, reasoned, “I personally use a rider’s rating as a guideline. So if a rider has a really low rating, it kind of is [that] my antennas are up so I will pay attention to behaviors that may be of concern that another driver has marked.” For him, a rating below 4.4 stars was “a really low rating,” so he might be more attentive to their behavior “just to make sure they’re not sneaking alcohol or things like that.” Comments like this reveals an internalization of the standard embedded in ratings: a high rating is equated with a “good” passenger.

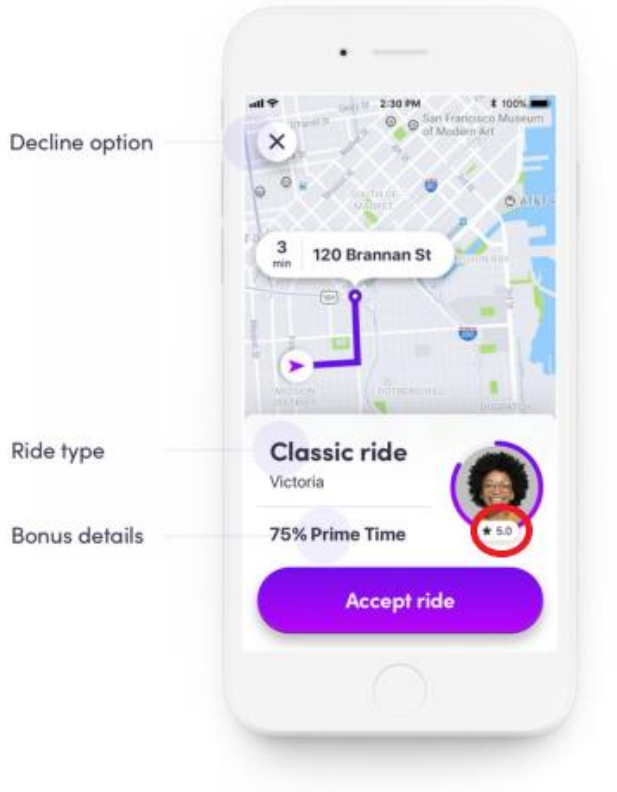
It was more common for my interviewees to dismiss riders’ ratings for two reasons. First, declining a ride simply means that drivers have to wait for another ride. Kimberly’s comment exemplified this point: “When you’re driving, if that car is empty, you’re wasting money. That car has to be full, so you want somebody in that seat the entire time … I’m not going to be picky about who I pick up.”

Figure 6-1. Screenshot of Uber's Ride Request on the Driver-facing App (Source: Hartanto & Attwell, 2019)



Note. I circled the rider's rating in red.

Figure 6-2. Screenshot of Lyft's Ride Request on the Driver-facing App (Source: Lyft, 2018)



Note. I circled the rider's rating in red.

Second, Uber and Lyft drivers only have 15 seconds to accept a ride request, meaning that they do not have much time to evaluate a rider based on her or his ratings. Crucially, not all the drivers could see riders' ratings *before* accepting a rider. Austin, for instance, told me that drivers in Chicago "don't get to see the ratings of passengers."[26] Uber does not disclose the information about where and why it decides to withdraw certain work-related information (i.e., riders' ratings) from drivers. Even if drivers could see riders' ratings, the interfaces of Uber (see Figure 6-1) and Lyft

[26] I also found a post about this issue on the subreddit for Uber drivers, but it is unclear whether Chicago is the only market where drivers cannot see this information.

(see Figure 6-2) matter in shaping drivers' attention. As Lyndsey explained:

> The way it's set up is you have this black window that shows up, and the wording [of riders' ratings] is so small that you really ... That doesn't really stick out. You see this black window here, and then you see the small white writing, and the biggest thing you see is accept, and then you accept it.

Alexis echoed this claim and suggested that "it flashes on the screen and we don't have enough time to look and study." Daniel, too, ignored riders' ratings because he only had a few seconds to respond to ride requests. Once he accepted a ride, he could not look at the ratings again. Hence, the temporal visibility of ratings may constrain how drivers could utilize such information for practical purposes.

TaskRabbit. Unlike the other two cases, TaskRabbit workers can communicate with their prospective clients before accepting and performing a task. They can also set their availability in advance and have longer time to respond to a client's request. Generally, there are two types of clients: *new* clients and *repeat* clients. While I will discuss how Taskers developed relationship with their clients in the next sub-section, the strategy I describe is mainly about the management of new clients.

All of the interviewees told me that they would ask a lot of questions about a task before accepting it to make sure that they and their clients would have the same expectations about the length and details about the task. For example, Scarlett was 71 years old and mainly performed the tasks of delivery and shopping. Before each task, she would ask questions about the delivery and pickup addresses, time expectations, the delivery items and their size, and whether any stairs would be involved. She

learned to ask all these questions after receiving a negative review because the client had hired her to organize for moving, packing, and unpacking delivery items, which required her to stand for long periods at work. She could not perform this task due to physical limitations; therefore, she had to clarify all of her limitations and understand her clients' expectations as clear as possible. As she explained, "more often than not, the clients are not very clear with what they want, and sometimes it's because they're poor communicators and sometimes it's because they're not clear themselves." During this process, she could also "profile" her clients before confirming the task.

Likewise, Easton focused on the pre-hiring process. He usually began by looking at clients' initial task explanation and figuring out whether they were "demanding" during the chat. As he offered,

> If they're very demanding or if they know how long something should take even though they're hiring someone to do it, that's when I say, "Yeah, this is not the person I want to work with." Because from my experience, when you get to the site, they're constantly trying to pressure you or check up on you … Especially if you're doing a job that you need to concentrate on. So I try to look for clients who overestimate the project length. Like for example, "Hey, can you put up a TV? I think it's going to take like two hours or something or one to two hours." They're being reasonable.

Put another way, this process of filtering helped Easton to look for "understanding and flexible clients" who were also likely to leave positive reviews and ratings.

This process of communicating expectations must take place via TaskRabbit's chat, even though clients may call Taskers to discuss their tasks. All of the

interviewees maintained that using the chat function was to document and protect themselves against unreasonable clients after completing tasks. As Neyland stated,

> Actually in each step with the client, it's very important to have the expectations set pretty clearly. So that if anything does go wrong or if they try to ask more or cause a dispute in regard to the job, then it's very clear in writing that nothing's going on.

He mainly performed moving tasks which might involve a second Tasker. As many clients did not make it clear, he had to make sure whether there would be another Tasker and what their responsibilities were. This is also to avoid miscommunication with the other Tasker because they would meet one another during service interactions. Kooper, too, shared:

> I've made sure to document everything in the chats … Everything that I talk about on the task and do [and put] it the in the chat. It [i.e., the chat] is a document piece in writing. If anyone has tried to or will try in the future to not 10 years say there's something wrong, we clearly discussed it in the chat.

The documentation was crucial for Taskers to dispute unfair ratings and file complaints to TaskRabbit. Additionally, some clients preferred phone calls over TaskRabbit chat. In this case, Audrey would write a summary of the call and ask clients to confirm it; otherwise, she would not show up to work for a task. For tasks like home repairs and furniture assembly, Taskers had to know what tools their clients could provide. As Murdock mentioned, some clients might be willing to cover the expenses of purchasing a tool if both of them did not have the tool. The chat allowed Taskers to document such additional expenses. Likewise, Kooper and Zoey both

encountered situations where their clients had payment issues such as declined credit card transactions. The documentation eventually helped them to show TaskRabbit that they did accept and perform the tasks in order to receive payment from the platform.

Indeed, the pre-hiring process is an essential way for both Taskers and their prospective clients to clearly communicate their expectations. It is likely that either Taskers or clients find that one another are not the best fit for the tasks. Cancelling tasks, however, would lower Taskers' acceptance rates. Murdock, for instance, did not want to accept tasks that were deemed "sketchy" when the initial task explanation did not provide much information. He reasoned, "the client can just use anybody's credit card and debit card and just sign up and hire somebody," compared to Taskers who need to pass background check. In response to "sketchy" requests, he would simply tell clients that he might not be "a good fit" and asked them to cancel the tasks. As he said, "If the client cancels, there is no risk on my half or my metrics. If I cancel, there's a risk of me losing that acceptance rates." Audrey, meanwhile, would directly inform clients about the negative consequences of cancelling a task on her end, if her clients requested her to do so. In doing so, she could avoid cancelling tasks to protect her acceptance rates.

Before discussing the practices of delivery workers, I would like to reflect on why the interviewees did not seem to focus too much on their online profiles. One of the key distinctions between TaskRabbit and the other two cases is that the former allows clients to look at the profiles of Taskers and hire a particular Tasker. Therefore, I initially expected that Taskers would engage in self-branding practices to showcase their expertise and maximize their earnings. However, the majority of the interviewees

(8 of 11) did not regularly update their online profiles except for their rates and availability. Easton, Maverick and Kooper were the only three interviewees who updated their online profiles to include the skills they had recently acquired and their work experiences on TaskRabbit, but they did so on a monthly basis. Easton, for example, updated his profile to include photos of his handyman services once a month. He told me that almost one third of his clients hired him because these photos showcased his expertise. By contrast, Maverick and Kooper mainly wanted to make sure that the description of their skills and experiences were accurate; therefore, they updated the profiles every three (Maverick) or six months (Kooper).

There are two reasons that explain why they did not engage in self-branding activities. First, the interviewees still got hired, even though they rarely updated their online profiles. Second, they already posted a lot of information—including the photos of themselves and their work as well as their tools, skills, and experiences—when signing up their Tasker profiles. Kuzma and Renee, for instance, spent a lot of effort in the online profile when they had started working on TaskRabbit. They wrote a long description for each category (e.g., landscaping and cleaning experiences) and they remained largely the same over time. Furthermore, it is difficult for them to demonstrate their expertise in delivery, moving, and personal assistant services. Highlighting the differences between his profiles on TaskRabbit and Upwork, Neyland noted,

> You have to prove yourself more on Upwork or any type of technological abilities … I would say I've updated my Upwork profile more because … it's pretty easy within a year to get exponentially better at programming while it's

difficult to get exponentially better at moving a couch.

His comment is indicative of how different types of platform-based work might influence the management of online profiles.

In a nutshell, the strategies of Taskers were similar to that of ride-hailing drivers, which both focused on filtering "difficult" customers before service interactions. Yet, the platform's chat function also enables Taskers to exercise their agency by documenting their interactions with customers.

Delivery Platforms. The delivery drivers were selective about the orders they accepted, but they did so largely because of earnings rather than customer ratings. Sean, for instance, shared, "DoorDash hides what the final payout is going to be for any order over a certain amount … If you get an order that is $8.50, you know that is going to be order that DoorDash is hiding the tip from you." Therefore, he tended to accept the $8.5 orders or above. Lorenzo and Connor deployed a similar strategy to accept orders. Other interviewees drew attention to the amount of customer tips as well as the restaurants. When Shawn started on DoorDash, he basically accepted all the delivery orders, but he soon found that some orders (e.g., driving nine miles for $3) were not worth his time. Therefore, he only accepted orders above $6 and customers who tipped. He, meanwhile, noted that customers who did not tip tended to give him a low rating. Collin added that he would look at which restaurant the customers ordered. If the restaurant was too far away or tended to have a long wait time, he would decline the order.

Akin to ride-hailing platforms, delivery drivers only have several seconds to accept an order. But compared to ride-hailing platforms, delivery drivers have even

less information about their customers before accepting orders because DoorDash and Instacart do not allow drivers to rate their customers. This might explain why the interviewees had limited ability to filter their customers to protect their metrics on delivery platforms.

Managing Customers During and After Service Interactions

Almost all interviewees, regardless of the platforms they worked, considered that being a "people person" was a required skill for them to navigate their work. It means that gig workers need to be patient and willing to communicate with their clients. As I illustrate below, my interviewees attempted to manage their customers' expectations and even their rating practices during and after service interactions. Taskers and delivery drivers, meanwhile, learned to document their services to protect themselves against difficult customers.

Ride-hailing Platforms. Ride-hailing drivers endeavored to allocate resources and perform emotional labor during service interactions and rate their "difficult" riders low after service interactions. First, the allocation of resources concerns the additional amenities the interviewees offered to their riders. Eddy, for example, offered Apple and Android chargers in his vehicle. Timothy, too, offered "extra amenities such as tissues, charger cables, the ability to play their music in my car, water bottles, gum, mints, and extra things like that." Clayton also gave out water and allowed his riders to use his tablet to watch videos and play video games. He tellingly added, "To me, the water is fundamentally setting a mindset for the customer of the difference between, 'You're just a transportation service,' and then, 'You're a more elite service.'" Alexis, meanwhile, attributed her positive ratings to the offering of additional amenities and

shared,

> I have Wi-Fi in my car. If you know my ratings, I have a lot of compliments on mine, and there's a reason for that. When people get in the back of my car, I have gum, candy. I have chocolates, even. I provide water. I have magazines. I have an umbrella for the elderly people.

While offering additional amenities may be ride-hailing platforms' intended outcomes of implementing rating systems, this strategy helped my interviewees to showcase their customer services.

Nevertheless, the interviewees recognized the additional expenses when they offered amenities to their riders. Noah, for example, shared, "What the customers don't realize with the charging the phone is that charge creates drain on the battery in my car. I have a $300 battery inside my vehicle." Despite the additional expenses, Noah continued to offer amenities to get good ratings. Lyndsey, by contrast, initially had offered such additional amenities, but then decided not to do so. As she explained, "In the beginning, I used to offer candy and water, but that began to become an expense … A lot of them, they won't finish the entire bottle, and they won't take it with them. So then that's added trash." For her, offering additional amenities created not only monetary costs but also manual labor to clean the vehicle.

The second strategy for managing customers was to perform emotional labor. This is analogous to the findings of extant studies (Raval & Dourish, 2016; Rosenblat, 2018a; Rosenblat & Stark, 2016). To begin, the interviewees had to make sure they had a correct destination address when their riders got into the car. Although riders had already input their destination when requesting the ride, the interviewees wanted

to check with them to correct for when customers entered the wrong address. In this process, Lyndsey would also ask her riders whether there was a particular direction they wanted to, or if they simply wanted her to follow the GPS. Then, she would inform riders about the estimated time arrival based on the GPS and check if they wanted to turn on the air conditioner or need anything.

During their expositions, interviewees recurrently used the terms—being "professional" and offering "good services"—to describe their service interactions to maintain a high rating. Charles, for example, reasoned, "What it takes to make a customer happy is good service." Good service, he continued, means "speaking to them professionally and providing them with everything that they need to get to their destination, and getting them there safely." Other interviewees attempted to spread "positivity" to their passengers. Alexis never made any complaints about Uber during service encounters. She explained,

> Because they [passengers] do not want to get into an Uber car and hear me complain about my job. If you got in the car and they would hear, 'Ah, I hate my job, blah, blah, blah,' if they heard me complain, they would rate me lower … I want to them to give me the best rating possible.

Her comment is indicative of the self-management of emotional displays during service encounters. She could not show her "authentic" feelings about Uber to her riders in order to maintain a good rating.

While Charles and Alexis focused on the ways of interacting with their customers, a small subset of the interviewees attempted to "game" or manipulate the rating system. I use the term "gaming" here because Uber and Lyft prohibit drivers

from asking for a 5-star rating. Figure 7 is a sign about ratings in one of my interviewees' vehicle, which is telling of this gaming practice. The sign states, "TIP NOT REQUIRED BUT GREATLY APPRECATE. DON'T FORGET TO RATE 5 STARS." The sign also asks riders to "enjoy the ride" and let the driver know if they need anything, which served as a signal to convey friendliness and remind riders to give the driver a 5-star rating.

Figure 7. A Sign about Uber's Ratings in An Interviewee's Vehicle

After service interactions, Uber and Lyft allows drivers to rate their riders. As previously discussed, the interviewees believed that the platforms would not punish riders who had a low rating. Yet, Lincoln, Rodrigo, Dominic, and Clayton knew that if they rated their riders 3-stars or below, Uber and Lyft would not match this rider with them in the future. While Clayton considered this was a loss of employment opportunities, other interviewees suggested that this would allow them to avoid having the difficult customers in the future. Lincoln, for instance, rated a rider 3-stars because

the rider slammed the door of his car. Rodrigo, meanwhile, articulated his evaluative standard. As he explained,

> The only time that I will rate a customer a one is if I am going to flag account, meaning the customer, meaning the customer tried to sneak alcoholic beverages into the vehicle … A two star is generally everything went wrong in that call, but you didn't break any Uber rules. So maybe you were rude, [or] you weren't courteous … That's when you'll either get a two or three depending on how severe it was … Usually, a four star is something like where everything was good, but you know, the time period for you to get to the car was six or seven minutes so I had to pass complete … Five star obviously, that kind of speaks for itself. That means that the customer was really good and it's not necessarily that they had to have a great interaction or anything because some customers, they just want to be left alone.

When Rodrigo articulated this evaluative standard, he was aware that Uber and Lyft might not punish riders because of his ratings. But he still wanted to rate the interactions based on his actual experience, in part because he would also reference riders' ratings when picking up riders. The articulation of evaluative standard might not directly affect riders' eligibility on the platforms, but it allowed Rodrigo to gain a sense of autonomy in this process.

TaskRabbit. In line with the earlier discussion about Taskers' practices before service interactions, interviewees attempted to manage their clients' expectations, document their service outputs, and develop long-term relationships with their clients.

Recalling Kooper's quote I mentioned in the introductory chapter: "over-

communication" enabled him to maintain his ratings. On the day he performed any tasks, he would double-check with his clients about whether they would be at the apartment and the address, and more specifically, "where to park or how to access the building, especially if it's a condominium or a private community." This was a common practice for the Taskers I interviewed. As clients could book a task up to 15 days in advance, this strategy helped to remind clients that they would be coming for performing the task. Kooper, meanwhile, tried to get to the place earlier to avoid taking too long to access to the property because it could delay the tasks, and therefore, lower his ratings. This may seem subtle, but he had received negative reviews because he was unable to discuss with the client in person. During and after service interactions, Kooper would take photos of his work and uploaded them in the chat. As he shared, "The client can see the progress of your work. This has helped immensely with getting higher ratings." Zoey, meanwhile, tried to make sure that her clients were at places where she performed the tasks. If the client would not be there, she would document the details about the tasks via TaskRabbit's chat for the purpose of documentation. As she put it, "I try to make sure everything is in the chat, so if something goes wrong, I have proof." As previously discussed, such documentation allows Taskers to dispute unfair ratings.

Although the interviewees attempted to gauge the scope of tasks before confirming them, clients' description might be different from the actual situation. Kuzma offered an example:

> Sometimes, the clients will hire me for an hour and a half or for a two-hour job. And I'll get there, and I'll see four bookshelves, and a bed, and two

> dressers, and I'm like, "Okay. That's like a four-hour job." So once I realize that, I'll discuss that with the client. I'll say, "Hey. Looks like this will probably take me a little bit longer than two hours. Are you okay with it taking about four hours if I'm here?" ... But obviously, for my schedule, that pushes things back a little bit. So once I realize that happens, then I'll try to schedule my clients with a little bit of a buffer between, but I also want to maximize how many jobs I'm working in a day … I don't want to be standing around for an hour doing nothing.

Kuzma told me that his clients were usually "very understanding" in this situation, but it was vital for him to manage the expectation about the length of the tasks immediately after he arrived. The interview quote also pointed out that this situation might delay other tasks he needed to perform on the same day. Hence, while he tried to not schedule any back-to-back appointments, he would tell clients whether he had any tasks before their appointment so as to manage their time expectations. Maverick, meanwhile, typically scheduled at least half an hour in between tasks "to account for any possible contingencies." If he completed a task earlier, he would inform the next client to see whether they wanted him to start the task before the scheduled time. He found this scheduling strategy useful to make his clients happy.

Similar to the case of ride-hailing drivers, Scarlett, Easton, and Kuzma explicitly asked their clients to rate them after completing tasks. None of them had explicitly asked for a positive rating; instead, they only reminded clients to review their performance. Scarlett, for example, described, "My closing phrase, which I haven't changed for a couple of years probably, I do ask people and I say, 'If you want

to help me out, please leave a review' … I do have that little blurb when I close out an invoice asking for a review if they have the time." Yet, she added that not every client would eventually rate her even after the reminder. Easton, similarly, would ask clients to rate him because as he put this, "it sort of pressures them to not put a bad rating in." Kuzma, by contrast, determined whether he would ask for a review, depending on how his interactions with clients went. When he believed that the interactions went well, he would remind the clients to tip and review him when sending an invoice to them. He figured out that about 75% to 80% of his clients would review him positively after using this strategy. However, he would avoid reminding clients to review him when he thought that he "had a weird relationship with the client," or that the client was not responsive. As he reasoned, "I just don't say anything about the review. So I guess, in a way, if I'm not reminding people about a review, then I might be avoiding people giving me bad reviews."

What distinguishes Taskers from the other types of gig workers is that the former could develop long-term relationships with clients. In the case of ride-hailing platforms, only Noah told me that he could potentially solicit riders as his private clients during Uber Select rides (i.e., a luxury ride-hailing option on Uber), but he chose not to do so because he could earn more money on Uber. Indeed, ride-hailing platforms may deactivate drivers who have contacts with riders outside of the platforms. By contrast, TaskRabbit encourages Taskers to do so, but such transactions must take place on the platform. The repeat clients usually were more flexible and gave Taskers positive ratings. Eight of the interviewees had repeat clients. As TaskRabbit offers discount to repeat clients, repeat clients might simply hire the

interviewees again on the platform. The only pitfall of having repeat clients on the platform is that Taskers could not automatically increase their rates for repeat clients. As Audrey described, “Whatever they hired you for, they will keep that rate for as long as you worked for them.”

Accordingly, four of the interviewees tried to develop long-term relationships with the clients *outside* of TaskRabbit. When asked about his strategies for establishing such relationships, Jeffery shared,

> I have a business card, and I put 10 or 20 of them in my wallet, and after each job, I would give them the business card and say, “If you want to hire me for anything else, please text me or call me at this number, and we’ll schedule it right there. No need to hire me through TaskRabbit.”

In doing so, repeat clients did not need to pay the 15% service fee to TaskRabbit. If the clients agreed to do so, he would charge them in cash or Venmo. Audrey, similarly, had long-term clients outside of the platform, but she would only accept requests from the “trustworthy” clients and avoid discussing any of these transactions within the app. Likewise, Kuzma pointed out that having transactions with clients outside the platform is a “gray area.” As he explained, “I’m not supposed to take my clients off the app, but if my clients request that they work for me in the future, I’m not going to say no, and obviously I personally prefer it that way.” Therefore, he would only give his business card to clients who explicitly asked him to do so. Importantly, when such transactions occur outside of TaskRabbit, Taskers no longer need to concern about their metrics and can negotiate their hourly rates with the clients.

Delivery Platforms. Compared to the other two cases, delivery drivers had limited interactions with their customers, due to contactless delivery during the COVID-19 pandemic. Collin, for example, said, “Many customers now only want you to leave the order in front of their door, they don’t want you to wait for them to give it to them or anything.” Ronnie, too, shared, “Especially with COVID, there’s very limited customer interaction.” But when asked about his strategies to get a good rating, Ronnie emphasized the importance of communication:

> You want to provide the utmost best customer service. Usually what I do is, when I receive the order, and let’s say for instance, when I arrive at the restaurant and the order’s not ready, I’ll message the customer and I’ll tell them, “Hi, this is Ronne. I am your DoorDasher, I’m here at McDonald’s picking up your order however, your order’s not ready at this time. Once I do receive it, I’ll let you know and I’ll get that to you as soon as possible.”

Roger, meanwhile, added that, “If they [i.e., customers] want me to give it [i.e., the order] to them, I’ll make sure to smile and say their name as I give it to them.” Both Ronnie and Roger wanted to convey their friendliness to their customers during service interactions, even though such interactions were limited.

As such, contactless delivery simplifies the delivery process, but it meanwhile creates another form of work-related uncertainty, as exemplified by Connor’s comment: “Negatively, you have to trust that the customer will take their order off their own doorstep, which sometimes can lead to the order disappearing altogether or them claiming they never received the order.” Connor, meanwhile, added, “there’s a lot of scamming going on now with customers.” As I discussed in Chapter 5, Clara

had received a negative rating because a delivery order was claimed to be missing. In order to reduce this form of uncertainty, delivery drivers learned to document photographic evidence of delivery. Indeed, DoorDash, Instacart, and Uber Eats also required delivery drivers to do so. But this strategy helped delivery drivers to protect themselves. Lorenzo, for example, shared, "I take a lot of pictures of things that I've dropped off just as protection because I've had people report that I haven't delivered things when I have." Sean echoed this and shared, the "picture proof" resulted in "a lot less issues with people claiming that I didn't deliver the food when I actually did."

Overall, the pandemic has significantly reduced the amount of service interactions between delivery drivers and customers. In other words, most of the interviewees did not seem to feel the necessity to manage their service interactions.

The Emergence of Worker-to-Worker Communities

Taking a step back, one might wonder how workers learned to manage their customers and metrics as forms of everyday resistance, especially when considering the atomized nature of platform work. I argue that worker-to-worker communities emerge as key social spaces for workers to seek advice and socialize with one another. Due to my sampling strategies, it is not surprising that the majority of the interviewees (42 of 50) had consumed content via worker-to-worker online forums or social media groups. Only a small subset of the interviewees (9 of 50) actively created and shared work-related information in the case of ride-hailing platforms. I consider these content creators as "driver/bloggers" to indicate their dual identity: they were driving for Uber, while also creating content about work-related tips and tricks on their blogs or YouTube channels. Before discussing these driver/bloggers, I will first explain how

and why the interviewees participated in worker-to-worker communities to learn to navigate the opaque labor platforms. Then, I turn to discuss the practices of driver/bloggers who might be considered the active members of these communities. The goal of this section is to demonstrate the potential of these online communities for individual and collective empowerment.

How and Why Participating in Worker-to-Worker Communities?

Before discussing the aspirations of driver/bloggers, it seems important to address how and why the interviewees came across these networked communities. When asked about where they got advice about their platform-based work, the interviewees commonly referenced online forums (e.g., Reddit), social media, and blogs as their primary information sources. As previously discussed, this is in part because platform owners do not provide any organizational training to workers and in part because platform-based help desks did not seem to understand workers' daily activities. The interviewees started taking part in these worker-to-worker communities because they had questions about taxes, changes in platform policies, and tips and tricks for maximizing their earnings and navigating labor platforms. Collin, for example, browsed through the subreddit for Uber Eats drivers when he ever needed advice because as he put this, "Uber customer service isn't really great, so I honestly avoid them whenever possible." As Uber classifies drivers as users of its technological service rather than employees, the company encourages drivers to seek "help" or contact its "driver customer service" when drivers have any questions. Murdock, meanwhile, described, "When I'm browsing through Reddit, I look for advice on how to improve my work for Tasks and adding categories." Derek also checked the

subreddit for Instacart shoppers to “get experience from others” every day.

Roger, similarly, considered the subreddits for delivery drivers as a place for him to look for solutions to a specific problem or question such as “how certain ratings affect my ability to get orders.” Roger learned from Reddit’s discussions that “acceptance rate means nothing” to Uber Eats and DoorDash drivers because the two platforms would not deactivate their employment. As he stated, this “was probably the most valuable thing for me” because “if the acceptance doesn’t ruin it [i.e., the continuity of employment], then I can choose what I want to do.” Connor, too, learned from online communities that “the real way to make money on DoorDash was to not accept everything they throw at you.” He considered Reddit as a “perfect” place for him to have “firsthand discussions with other drivers.” Their reflections indicate the potential for workers to be empowered through information seeking and sharing practices.

Crucially, these worker-to-worker communities provide social spaces for gig workers to interact with their peers and to feel connected. When the interviewees such as Eddy, Adrian, Melvin, and Nelson had work-related questions, they would look at these online communities because they wanted to know their peers’ opinions. As Nelson put this,

> On the subreddit, it gives you a more real look. Whatever DoorDash is going to say, it's going to be putting them in a good light in trying to make certain things seem a certain way according to them, while getting information from actual users of the app will give them unbiased type of look at it.

Similar to Connor’s reflection above, Nelson took part in worker-to-worker online

communities because other participants were “actual users of the app” who could provide firsthand and realistic perspectives of the platform.

To be sure, as gig workers become familiar with the worker-facing app, they might not feel the need to continuously look for advice within these online communities. Jeremy, for instance, said, “They were very useful when I was starting out. Now that I have more experience, I don’t feel the need to use them as much. But I do think that they are a valuable resource for people who need them.” He tellingly explained the importance of these online communities:

> One of the things about not just being an Uber driver, but also being a driver in general … is that you’re not really interacting with other drivers per se, unless you do it through the internet … Using Grubhub as an example. There’s this one Facebook group for Grubhub drivers that’s dedicated to Grubhub drivers and they’ll share stories, they’ll have jokes, they’ll go socialize on the internet. It’s basically the online water cooler space for other drivers. There’re Lyft and Uber driver Facebook groups that are similar to that.

As Uber and Lyft drivers, and more generally, gig workers do not have co-workers in a physical workplace, these communities were crucial for them to share stories and form social networks. This may explain why Melvin and Sofia followed the subreddit and meme pages for DoorDash drivers for “entertainment” purposes because these communities allowed them to know the positive side of their peers’ work experiences. When Scarlett reached 500 reviews in her 1,000 tasks, she posted this accomplishment in both official and unofficial Facebook groups for Taskers. As she recalled, “I posted that just bragging, just saying how proud I was in myself. And you get a lot of positive

feedback in that time."

Interestingly, TaskRabbit creates an official Facebook group for "Elite Taskers." A TaskRabbit's community manager moderates this group. The community guideline for this Facebook group encourages Taskers to "help each other," "refer great Taskers," "use all your resources," "keep it professional," "update the community," and share "best practices" and "Tasking stories" (TaskRabbit, n.d.-b). The moderator may also remove posts that disclose clients' information and share "inaccurate information about the platform that could misinform Community members." In practice, the interviewees who were part of this Facebook group were aware of TaskRabbit's moderator, but they added that this moderator never responded to any posts in this Facebook group.

Indeed, the interviewees told me that Taskers shared complaints about TaskRabbit in the official Facebook group. Kuzma offered some examples of the posts in this group: "A lot of people [were] complaining about the jobs that they've had that they just didn't like, or problems with the app that just no one ever really help us with." He initially had thought that the moderator from TaskRabbit would look at the Facebook group and fix the problems, but in fact they did not. Jeffery, too, shared, "TaskRabbit doesn't actually watch what we're saying in TaskRabbit. Thank goodness, because we talk about ways to cheat the system." By "cheating the system," he referred to the practices of forming long-term relationships with clients and asking these clients to use their service *outside* of the platform. Scarlett, by contrast, recalled that TaskRabbit's moderator participated more in the past, but now "it's pretty rare." This official Facebook group, for my interviewees, eventually became a virtual place

for them to get to know other Taskers' stories. At the same time, almost all of them took part in unofficial Facebook groups for Taskers for information-seeking purposes. Renee and Scarlett noted that the unofficial groups had more members and posts than the official one, but if they ever had work-related questions, they would go to both official and unofficial groups to seek help.

Taken together, the interviewees primarily consumed content and sought work-related information within these worker-to-worker communities. These communities emerge as online occupational communities (Schwartz, 2018) where workers are socialized to adapt to navigate labor platforms.

Uber Driver/Bloggers: A New Cast of Experts in the Gig Economy

What distinguishes what I call "driver/bloggers" from the other interviewees is that the former actively created content about Uber and positioned their expertise regarding the inner workings of the platform. Taking a step back, there exists a new cast of self-proclaimed experts offering "how-to-succeed" resources aimed at coaching and inspiring gig workers, as evidenced by self-help books such as *The Gig Economy: The Complete Guide to Getting Better Work, Taking More Time Off, and Financing the Life You Want* (Mulcahy, 2016) and *Thriving in the Gig Economy: How to Capitalize and Compete in the New World of Work* (McGovern, 2017). The driver/bloggers I discuss in this sub-section can be considered a group of these experts. By April 2018, the most subscribed Uber-related YouTube channels had about 64,000 subscribers, but most of the channels only had several hundred or a few thousand subscribers. The driver/bloggers (n = 9) I interviewed had a network of subscribers ranging from 100 to more than 20,000. Regardless of the number of their

subscribers, they actively created videos on YouTube on a weekly or bi-weekly basis. Below, I focus on their vision of collective empowerment and individual aspirations.

All of the driver/bloggers I interviewed had a vision of empowering their audiences through the distribution of actionable information about Uber's labor platforms. Indeed, they started their YouTube channels or blogs because they could not find the *relevant* information when they had worked as an Uber driver. Here, relevance refers to a particular angle to approach the information from drivers' perspective. Austin, for instance, explained, "The reason I started was because I didn't see anyone else talking about realistic expectations about what you could achieve monetarily with rideshare." In a media interview, Harry Campbell, who operates "The Rideshare Guy," attributed the popularity of his blog to Uber's and Lyft's poor customer service for drivers: "I sometimes joke that if Uber and Lyft were better about customer service I wouldn't have a job" (Matsakis, 2016). Inherent in their reflections is a sense of information asymmetries between the company and drivers as well as among drivers themselves.

We can also observe the vision of collective empowering by looking at driver/bloggers' self-descriptions and videos on YouTube. One driver/blogger articulated his channel as a site to "inform new, current or future drivers on how to drive for Uber and Lyft successfully," whereas another claimed to "share tips and information to help drivers make more money" in the self-description. Similarly, reflecting what he aspired to achieve through the YouTube channel, one driver/blogger stated in his video, "I want to set up a network of experienced and professional drivers that are willing to help out other drivers and have this be a team, a

community." The narrative of empowerment is grounded in a premise that "no one is an expert when first start doing this [Uber]," as Austin put it. As drivers become acquainted with Uber's platform and policies after seeking information online, Austin continued, "some of these new drivers are becoming experts" and they "could financially improve their lives." Or, as Cohen explained, "we become an expert in driving for Uber, if we do it enough."

YouTube offers a networked space for driver/bloggers to express themselves and create a sense of community. Reflecting on her experience of getting started, Alexis offered, "I wanted to voice my opinion … I love the community of people that are YouTubers. We have a common bond, because not only do we do rideshare, but we do YouTube." Lyndsey, meanwhile, envisioned her channel as "a resource channel." As she explained,

> I want it to be a channel that Uber drivers and Lyft drivers from all over who have access to the live [i.e., her live stream on YouTube], that they have a place that they can come to immediately and get someone to ask and tell them, okay do this, do that. Or this is what happened when I did that and give different suggestions. Not just mine, they are other Uber and Lyft drivers that are part of my live.

While Alexis' comment articulated the connection among the driver/bloggers, Lyndsey pointed to how YouTube enabled her to build communities with her audiences who were also Uber and Lyft drivers.

Alongside the vision of collective empowerment, the interviewees articulated their individual aspirations of developing dual "gig" careers as Uber drivers and

bloggers. Consider, for example, the video course "Maximum Ridesharing Profits" (US$97) offered by "The Rideshare Guy" (i.e., Harry Campbell). He regularly interviewed drivers for his blog and YouTube channel. In a YouTube video about his course, he described that the course was based on "everything we have learned from interviewing drivers." As he put this,

> You start from our course and you'll be making money from your very first day … Even if you're an advanced driver and you think you know everything, I guarantee you will learn at least one thing in our course. And if you don't, send me an e-mail and we'll talk. And I'll take that as a challenge to figure something out.

Another example is the "Rideshare Driving School" founded by the "Rideshare Professor." He offered a series of courses about ride-hailing platforms (US$69 for three series of the courses). In the video about the "Rideshare Driving School," he stated that the courses would share his "success formula" to help people to increase their earnings. The content of the courses included not only the tactics to navigate Uber's sociotechnical platform, but also skills in curating a brand on social media to promote referral codes and side businesses. Indeed, asking subscribers to use their referral codes was a common way for driver/bloggers to make money. Oftentimes, such referral codes were not exclusive to Uber or Lyft, but also other labor platforms such as Instacart and DoorDash. When people use driver/bloggers' referral codes to sign up as workers on labor platforms, both of them can get a referral bonus.

Among the driver/bloggers I interviewed, Cohen was the only one who had successfully monetized his expertise in Uber through advertisements and referral

codes. He still occasionally drove for Uber and Lyft and hired drivers to write about their stories and work-related tactics on his blogs. This is because he considered that his audiences wanted to "learn from someone who [is] struggling with this [i.e., Uber]." The rest of the interviewees were still struggling to make a profit from their channels. When asked about how he envisioned the future of his channel, Charles shared, "[R]ight now I make way more money with Uber than I do with YouTube creating. I'll be able to create content and feel like, 'Hey, wow. I can kind of equal this thing out. I can make just as much with YouTube as I can with driving for Uber Eats.'" Clayton, too, aspired to get more subscribers and figure out his personal brand in the future, but it also meant that he had to spend more time on creating content rather than driving. Yet, as the channel grew, Austin added, "It's becoming more challeng[ing] to just talk to everyone and get back to them." While interviewees like Austin initially did not expect their channels to grow and monetize their expertise, concerns about monetization eventually emerged, as they engaged in incessant invisible labor (Duffy, 2017) to make videos and cultivate social relationships with their subscribers.

In summary, the case of Uber driver/bloggers sheds light on what motivated drivers to create content to empower their peers and pursue individual aspirations. While Uber drivers cannot communicate with their peers on the platform, driver/bloggers' practices cultivate a sense of community on another platform—YouTube. Although driver/bloggers' content creation and the interviewees' information-sharing practices took place outside of labor platforms, they still had a potential to empower gig workers in at least two ways. First, as some of these work-to-

worker communities, particularly YouTube channels are open to the public, they can also have a potential to raise public awareness of the labor conditions in the gig economy by unmasking Uber's unfair practices (Ticona et al., 2018). Second, these online communities emerge as important social spaces for gig workers to interact with their peers and learn work-related strategies to navigate labor platforms. Workers could learn how to interpret their metrics and manage their customers from these networked communities.

Summary of the Chapter: The Implications of Everyday Resistance

In the first part of this chapter, I discussed the ways in which workers attempted to manage their customers before, during, and after service interactions, so as to protect their ratings in the three cases. Before service interactions, ride-hailing drivers and Taskers tried to filter "difficult" customers. While ride-hailing drivers might strategically avoid providing carpooling services and pay attention to riders' ratings, Taskers attempted to manage their clients' expectations about the tasks and consciously use TaskRabbit's chat function to document the pre-hiring process. During and after service interactions, ride-hailing drivers mainly offered additional amenities and performed emotional labor in exchange for a good rating. They might also develop their own evaluative standard and explicitly ask riders to rate them. Taskers, meanwhile, managed their clients' expectations, documented the task outputs, and developed long-term relationship with their clients. Compared to the other two cases, the pandemic has significantly reduced communication between delivery drivers and their customers. Therefore, I did not find much about how they attempted to strategically manage their customers to protect their ratings. But as I discussed in

Chapter 5, delivery drivers still experienced uncertainty around their metrics.

The management of customers reveals workers' strategic attempt to communicate with their customers to gain a sense of autonomy over platform-based labor processes. While platform owners might favor customers over workers through the implementation of customer ratings, workers could also find ways to protect their ratings. Many of the practices I documented might seem subtle and routinize service interactions, but they were essential for the interviewees to minimize the interactions with difficult customers and to ensure that customers would be satisfied with their services. I consider these practices as moments of everyday resistance, in part because the interviewees were keenly aware and critical of the disciplinary outcomes of metrics. It was difficult for them to completely ignore metrics; therefore, the management of customers became a viable way for them to routinize and control service interactions. Indeed, the practices of offering additional amenities, performing emotional labor, and managing customers' expectations might be in line with the intended outcomes of labor platforms. These practices, however, also enabled the interviewees to manage difficult customers who tended to disrupt service interactions and rate them low. The practices of filtering customers and documentation might, to a certain extent, allow workers to take advantage of socio-technical features of labor platforms without breaking the rules. Among all these practices, asking for reviews and developing long-term relationships with clients outside of TaskRabbit might break the rules of labor platforms. But at the same time, they were largely invisible from the platforms because platform owners could not monitor service interactions.

While the management of customers reveals workers' daily practices, the

second part of this chapter reflects on the potential for workers to seek work-related information and articulate visions of collective empowerment. The creation of information asymmetries is a central characteristic of algorithmic labor management (e.g., Rosenblat & Stark, 2016; Shapiro, 2018). Labor platforms are often assumed to be individualized and atomized workplaces because gig workers work independently and do not have any co-workers. As I have demonstrated, worker-to-worker online communities emerge as social spaces for workers to socialize themselves to adapt to platform-based work. This may also explain how and why my interviewees articulated their understandings of metrics, despite the opaque platform policies and socio-technical features. Even though workers became familiar with the platform's policies and socio-technical features, they still considered such communities as valuable resources. The case of Uber driver/bloggers has shown how experienced drivers might articulate their expertise and seek to empower their peers outside of Uber. Meanwhile, the driver/bloggers had an aspiration to garner visibility and monetize their expertise in the future. As driver/bloggers developed their dual careers as Uber drivers and YouTubers (or bloggers), they arguably engaged in what Duffy (2017) calls "aspirational labor," a form of mostly uncompensated labor with a hope for future economic opportunities. Taken together, the emergence of worker-to-worker online communities showcase the potential for empowering gig workers within collaborative social networks (see also Gray et al., 2016), while giving rise to a new cast of aspirational brand-builders (see also Rosenblat, 2018a).

Overall, this chapter has discussed how workers can manage "imperfect" metrics by coping with customers and reworking platform-based labor processes on

and off labor platforms. Of course, only a few workers would actively create content and aspire to develop a dual career. But workers who seek work-related information from these networked communities can familiarize themselves with platforms' policies and learn how to manage metrics that matter to their employment opportunities. In this vein, worker-to-worker communities can socialize newcomers to adapt to platforms' algorithmic management. This process of socialization is essential for workers to learn to engage with moments of resistance in their daily work practices. Although the management of customers are constrained by labor platforms, they help to reduce work-related uncertainty and allow workers to gain a sense of autonomy over metrics that are beyond their control.

CHAPTER 7

CONCLUSION

How do algorithmic metrics perform as technologies of control and resistance in gig workers' everyday work practices? Through a comparative case study of digitally-enabled service work, this dissertation has examined the power dynamics underpinning the management, interpretation, and uses of metrics—including customer ratings and system-generated behavioral measures—in platform-specific calculative spaces. Specifically, I have discussed how labor platforms—including ride-hailing platforms, TaskRabbit, and delivery platforms—deployed metrics in their labor processes for managing workers, while simultaneously framing metrics as a boon to the entire community of workers and customers. Second, I have explicated how metrics facilitated the tensions between the management *by* and *of* customers. Third, I have explored the similarities and differences across the three cases. The comparative analysis has demonstrated how and why certain kinds of metrics could facilitate algorithmic labor control through the production of uncertainty and anxiety. I have also discussed how workers attempted to manage metrics by strategically managing their customers and sharing work-related information.

In this chapter, I begin by summarizing the empirical findings. Then, I discuss the implications of the findings for understanding service triangle in the gig economy and the limits of metric power. Finally, I discuss the limitations of this project and suggest future research directions.

Summary of the Findings

Using a qualitative multi-case study approach, this project has examined how

digitally-enabled service workers encountered, interpreted, and managed metrics across three kinds of labor platforms—ride-hailing platforms (i.e., Uber and Lyft), a temp work platform (i.e., TaskRabbit), and delivery platforms (i.e., Uber Eats, DoorDash, and Instacart). Originally, I planned to study platform co-ops, but I had to drop this case due to the COVID-19 pandemic. Eventually, I included delivery platforms, which have experienced an explosive growth due to the pandemic, as the third case. The three cases represent distinct types of service work and visibility of platform-based metrics. In each case, I drew on the analysis of discursive texts related to these labor platforms and in-depth interviews with workers. Motivated by understanding metrics as labor control and resistance, this project asked three research questions in the introductory chapter:

1. How do platforms use metrics to govern their distributed workforces?
2. How do workers interpret and manage metrics as part of their daily service interactions and work practices?
3. What are the factors that shape workers' interpretation and uses of metrics across labor platforms?

In response to the first research question, I explicated the role of metrics in platform-based labor processes. By design, metrics become interwoven with platform-based rewards and sanctions. Platform owners require workers to have a certain level of customer ratings to qualify for rewards programs, which might further affect their access to work-related information and employment opportunities. Except TaskRabbit, the other platforms I studied rely on metrics, particularly customer ratings to determine workers' continuity of employment. While customer ratings become

ubiquitous in different service sectors, labor platforms do not simply consider ratings as reference points, but rather use them to determine employability. Through the deployment of customer ratings, platform owners can monitor and discipline workers, while attributing the responsibility to customers. In this process, workers have limited communication with both platform owners and customers concerning the accuracy and fairness of ratings. Meanwhile, platform owners emphasize the importance of metrics for the entire community and equate high customer ratings with "good" and "professional" services. In other words, metrics entail both technical and normative control. More broadly, we must consider how metrics become normalized as meaningful proxies for performance and even reputational capital in our daily life.

The issue of labor control is interrelated to the second question. There are two types of metrics, namely, customer ratings and system-generated behavioral measures. I found that most of the interviewees expressed frustrations over customer ratings rather than behavioral measures, in part because the former directly affects their employability and in part because the former involves the production of uncertainty. In most cases, workers were not rated by their customers, which led to work-related uncertainty. A few negative ratings could potentially affect workers' average ratings, and thus their employment opportunities. This issue was more salient in Uber than Lyft because Lyft would automatically give drivers a 5-star rating when customers do not rate. Even when workers did get rated, I found that workers had insufficient actionable information to interpret why customers rated them low, and even if they speculated for plausible reasons, those factors were often out of their control. Indeed, workers drew attention to "difficult" customers who had unreasonable expectations on

their work. In the case of delivery drivers, restaurants could threaten workers' ability to maintain their customer ratings. Metrics, therefore, are "imperfect" systems for service management because they do not allow workers to "improve" their performance, even if workers want to do so. More importantly, the affective feelings and uncertainty provoked by ratings speak to a wider concern over the trilateral relationship between platform owners, workers, and customers. The deployment of customer ratings facilitates and rationalizes the management of workers by customers.

Workers, meanwhile, learned to manage their customers to maintain a sense of autonomy in their daily work practices. This involves the filtering of customers before service interactions, management of customers' expectations during and after service interactions, and gaming practices around customer ratings. The management of customers seemed to be a scripted outcome of platform owners, but it meanwhile helped workers to protect their ratings and routinize service interactions. Gaming practices—asking for 5-star ratings (ride-hailing platforms) and developing long-term relationships with clients outside of the platform (TaskRabbit)—reveal workers' strategic attempts to rework platform-based labor processes to exercise their agency. As such, workers' daily practices around metrics show the tensions between management by and of customers. Outside of the platforms, workers engaged in the practice of consuming and sharing work-related information within worker-to-worker online communities. Workers may work within a social network of their peers for empowerment, even though they are self-employed and often considered atomized labor.

When I started this project, I expected to find a story of divergence across

labor platforms. Indeed, I considered my case selection logic could maximize the variation of platforms which deploy metrics, with different degrees of visibility. For example, I expected that Taskers would engage in self-branding practices to promote their metrics and profiles, as opposed to ride-hailing drivers and delivery drivers whose metrics had limited visibility to future customers. Despite the variation of platform-based metrics, the interview findings show how workers similarly expressed anxiety and frustration over customer ratings. To be sure, the disciplinary outcomes of ratings varied. Among the three cases, TaskRabbit workers arguably had a higher degree of autonomy than the other cases for two reasons. First, their metrics were indirectly connected with the visibility of their online profiles, rather than the visibility of work-related information (ride-hailing platforms) as well as the continuity of employment (ride-hailing platforms and delivery platforms). Second, TaskRabbit had a relatively better dispute mechanism from the interviewees' experiences. Taskers could dispute unfair ratings as long as they documented their interactions with customers and the details about tasks. Yet, it is noteworthy that TaskRabbit, much like the other cases, offered a limited description of its dispute policy. Besides platform-based labor processes and dispute policies, I found that workers' economic dependence on the platform might affect how they would respond to metrics. Workers who were less economically dependent on the platform might be less worried about metrics, and therefore, subject to a lower level of algorithmic labor control. Future research may consider different dimensions of platform dependent and their relation to labor control and resistance (for an exemplar, see Schor, 2020).

To conclude this project, I would like to discuss two broader issues of

theoretical and practical significance. The first issue is about how metrics may reinforce the long-standing, unequal power-relations between managers, workers, and customers. The second one has to do with the limits of metric power; specifically, in what ways metrics are problematic and why we cannot simply fix these problems through technological design.

The Service Triangle in the Gig Economy

Central to the concept of service triangle is the three-way interest alliance in service workplaces (Leidner, 1993). The presence of customers is part of the inherent unpredictability of service interactions because customers are not required to follow the scripted service interactions. The interest of managers, workers, and customers would not only be in direct conflict, but also occasionally align with one another. But in practice, managers have long normalized the management by customers in service workplace (Bélanger & Edwards, 2013; Bishop & Hoel, 2008; Fuller & Smith, 1991; Leidner, 1993). Alongside the presence of customers, the case of delivery platforms adds a new layer of the three-way interest alliance, that is, restaurants. Platforms (or managers) and workers cannot control how restaurants process delivery orders, which might induce work-related uncertainty for workers to manage. Yet, the management of workers by customers remains a focal point in the three cases.

Specifically, I have shown how metrics, particularly customer ratings rationalize and reinforce customer-oriented rules. While gig workers are supposedly independent contractors—or "flexible" workers, as platform owners have promised—they are monitored and managed by customers. While the findings are similar to that in the literature around service work, there are three crucial differences in labor

processes and control through the lens of (in)visibility. As Beer (2016) aptly argues, "Metrics can be used to expose or conceal, to highlight or obfuscate, to illuminate or shade. In some instances visibility is to empower; in others, it is disempower" (p. 173).

First, what distinguishes digitally-enabled service workers from their predecessors is the temporal ordering of service interactions. Traditionally, service interactions take place when workers and customers meet one another, often in a physical location (e.g., taxi, customer's home, a restaurant, etc.). Now gig workers meet customers virtually via the app *before* service interactions. Theoretically, metrics, and more broadly, digital data can premediate both parties' expectations of service interactions. The key issue, here, is what data are visible to workers and customers at this point. Empirically, we can observe this by attending to (a) how Taskers anticipated their prospective clients would look at their profiles; and (b) how some ride-hailing drivers and Taskers might filter their customers before service interactions. In the case of TaskRabbit, making workers' metrics visible can be beneficial for customers to make hiring decisions, but it simultaneously leads workers to be anxious about the outcomes of evaluation. In the latter example, the temporal ordering of service interactions signals the potential for workers to claim a sense of autonomy. But ride-hailing drivers only have several seconds to look at customer ratings and other work-related information, whereas Taskers cannot rate their customers. Likewise, delivery drivers cannot evaluate their customers. If we accept that ratings represent proxies for trustworthiness and are valuable to the entire community, the absence of ratings about customers essentially means that workers

lack such informational signals when evaluating an order. This is not to argue that everything should be rated, but to question platform owners' unequal treatment of workers and customers. As such, platform owners withhold actionable information from workers, rendering workers difficult to evaluate what is worthy of their labor.

The second line of questioning has to do with the invisibility of feedback from customers and platform owners. This project contributes to understanding the affective power of metrics through the production of uncertainty and anxiety. The source of such affective power is the lack of qualitative feedback for workers to modify their performance. They can only rely on speculation. Previous studies have discussed how managers (a) monitor whether workers follow organizational scripts to interact with customers (Leidner, 1993; Woodcock, 2017) and (b) solicit customer feedback on workers' performance (Fuller & Smith, 1993). While managers are likely to favor customers over workers, workers at least know that they must follow the scripts and may be able to talk to their managers. But training and scripting are largely absent across labor platforms. Consistent with existing studies about algorithmic labor control in the gig economy (Gandini, 2019; Rosenblat, 2018a; Rosenblat & Stark, 2016; Stark & Levy, 2018), I argue that the deployment of customer ratings outsources managerial tasks and responsibility to customers. Hence, customers not only contribute to the unpredictability of service interactions, but also managerial outcomes in digitally-enabled service work.

The third distinction is about the (in)visible role of platform owners. As scholars have highlighted, platform owners attempt to distance themselves from any employment responsibilities (e.g., Rosenblat, 2018a; van Doorn, 2017). Seemingly, it

is not the platform owner that monitors and disciplines workers, but platform-based algorithms and customers who actively do so. The deployment of metrics, particularly customer ratings, reveals how platform owners shift managerial responsibility to customers. Nevertheless, platform owners remain a key position in the service triangle because they provide socio-technical infrastructures to connect workers and customers and to make service transactions possible. Through technological design and user agreements, platform owners retain the right to regulate the uses of metrics—and more broadly, the data generated by workers and customers—to manage their distributed workforces. On a discursive level, platform owners frame (virtual) rewards, such as high ratings, as recognition and metrics-based punishment as a means to protect the community. Metrics and rewards programs, meanwhile, define the meanings of "good" workers and behaviors by appealing to customer-oriented rules. Indeed, workers can dispute what they perceive to be unfair ratings, but platform owners determine how to handle it. Even though platform owners tend to take the side of customers in the three case studies, their decisions are still "legitimate" according to their "community guidelines." That said, while platform owners (or managers in the service triangle) seem to be invisible from processes of evaluation, they actively shape these processes to advance their interests.

Taken together, I contend that the deployment of metrics places workers in a vulnerable position in the three-way contest for control. Workers lack the ability to mobilize and manage metrics against the other two parties. They can still strategically manage their customers to gain a sense of autonomy, but such reactive practices are largely constrained by what information they have and how they can communicate

with customers via the platform. This project, therefore, contributes to understanding how metrics reinforce the authority of platform owners and customers, while disempowering workers in the service triangle. Moreover, I contend that evaluation is a key terrain of control and resistance because metrics bring together the three parties. While this project focuses on workers' perspectives, we can observe how metrics incite them to interpret their interactions with customers. Therefore, this project can contribute to our understanding of algorithmic labor control by approaching metrics and customers as sources of uncertainty.

The Limits of Metric Power

Throughout the research process, one question I kept returning to was: Are metrics inherently bad? Like many hard questions, the answer is: it depends on how metrics are being used, but perhaps more importantly, fixing the problems of metrics, particularly customer ratings are difficult and cannot tackle pressing issues about algorithmic labor control in the gig economy.

Empirically, we can observe several ways of framing and using metrics in service workplaces. Situating metrics into different platform-based labor processes, metrics become interwoven with the visibility of online profiles and the availability of work-related information as well as rewards and sanctions. Platform owners hold that metrics can promote accountability and mutual care. Customer ratings, ideally, can build trust among total strangers (Dellarocas, 2003) and create a "win-win-win" situation for service management. Ratings impose a normative standard for service encounters; for example, platform owners equate a high rating with "good" and "professional" services. Workers may also agree that ratings can be proxies for their

reputation and the service they provide. But the problem is that customer ratings do not always measure what is intended to measure (i.e., the quality of service), in part because many customers do not rate and in part because customers do not need to justify their ratings. The lack of actionable information about ratings, together with the disciplinary outcomes of ratings, produce uncertainty and anxiety.

Nevertheless, the discrepancy between intended purposes and actual uses of performance metrics—regardless of whether they are algorithmic or not—is a common issue when metrics become dominant modes of evaluation, as scholars of quantification have long argued (Espeland & Sauder, 2016; Hoskin, 1996; Muller, 2018; Strathern, 2000). In order to quantify and standardize, metrics must simplify the qualitative complexity of our social world. This process is always incomplete and entails subjective judgment about what counts and what does not count (Espeland & Stevens, 2008; Esposito & Stark, 2019; Fourcade, 2011). As Esposito and Stark (2019) argue, metrics as "the observation of others provides a reference point from which to observe oneself" (p. 19). This reference point is valuable but is also "simplistic, obscurantist, inaccurate, and subjective" (p. 10).

My point here is not to oppose all forms of (customer) evaluation of workers. Instead, metrics can be useful when they are merely reference points about the quality of service interactions rather than the determinant of workers' employment opportunities. This is more salient in the cases of ride-hailing and delivery platforms because workers have very limited ability to dispute unfair ratings and communicate with the companies. Assuming that ratings remain the determinant of employment opportunities, simply changing the design of ratings would be less likely to reduce

work-related uncertainty and anxiety (e.g., TaskRabbit versus Uber). Indeed, all the platforms in three cases have changed their design of ratings over time. Uber and Lyft, for example, may justify their implementation of "two-way" rating system and ratings protection by "improving user experience" and "learning from the community of drivers." Such changes do not resolve the disciplinary outcomes of ratings but become rhetorical justifications for platforms to reduce labor control to technological problems that can be fixed by technological means.

What is at stake here is not merely whether workers' performance should be evaluated, but the point that platform owners put too much emphasis on ratings and make it difficult for workers to dispute unfair ratings. Based on the findings, a possible first step for addressing the issues of metrics and labor control is to provide a meaningful explanation of ratings and dispute policies. Here, a meaningful explanation is not simply to make transparent the inner workings of algorithms, but to offer understandable explanations of automated data-driven decisions and provide data subjects with grounds to reverse and challenge the outcomes (Wachter et al., 2018). This is in line with The Fairwork Project's (2021) principle of fair management. Platforms should "provide due process for decision affecting workers." Such an appeal process should allow workers to "communicate with a human representative of the platform" and be available on the worker-facing app. Meanwhile, fair management is only one of the five principles, alongside fair pay, fair conditions, fair contracts, and fair representation (The Fairwork Project, 2021). This is also why fixing the problems of metrics cannot tackle all the issues surrounding algorithmic labor control. Even if metrics are no longer connected to the visibility of work-related information and

control the process of matching workers and customers, platforms could still rely on algorithmic direction and discipline to govern their distributed workforces.

Additionally, metrics not only facilitate algorithmic labor control but also shape how workers strategically manage customers and share work-related information to resist the impacts of metrics. The day-to-day management of customers may seem subtle, in part because their agency is constrained by platform-based algorithms (e.g., Anwar & Graham, 2020). But it is important to keep in mind that when such individualized tactics are widely shared through worker-to-worker online communities and adopted by workers, they can rework platform-based labor processes and even repurpose algorithms that rely upon data subjects' inputs. Consider, for example, the situation where a critical amount of Uber and Lyft drivers rate their "difficult" customers 3-star ratings or lower, meaning that they will no longer be matched with these riders. In this case, drivers could avoid providing service to difficult riders. These difficult customers might also have difficulties in requesting for a ride when drivers pay close attention to riders' ratings. Hence, the moments of resistance have the potential for individual and collective empowerment.

The key point that I would like to close with is that algorithmic labor control and resistance must be situated within a wider political economy. As van Dijck et al. (2018) aptly argue, we are in "a connective world where platforms have penetrated the heart of societies—affecting institutions, economic transactions, and social and cultural practices—hence forcing governments and states to adjust their legal and democratic structures" (p. 2). Put another way, platforms can contest regulations, but they are also governed by political institutions (Gorwa, 2019). The most recent

example is the “Prop 22” campaign in California, where labor platforms reportedly sent emails and in-app messages to persuade voters to support their sponsored bill (Hawkins, 2020b). Uber and Lyft required drivers and riders in California to tap the “confirm” button on the in-app messages about how the passing of Prop 22 could save workers’ employment opportunities, lower wait times, and avoid higher prices *before* accepting or requesting a ride. While Prop 22 was about gig workers’ employment status rather than platform-based metrics and algorithms, this political event indicates not only the political power of platforms but also the role of public in regulating such platforms. The employment status of gig workers is inseparable from algorithmic labor control. As in the case of Prop 22, labor organizations which oppose Prop 22 had to compete with labor platforms with financial capital and access to their large base of user groups. This raises further questions about how platform governance shape and reshape public values through popular mobilization.

Limitations

There are several limitations of this project. First, it is important to keep in mind that my interview sample was skewed to male gig workers and those who were likely to use worker-to-worker online communities. There are two possible reasons for the demographics of the sample. First, methodologically, I mainly recruited participants through worker-to-worker communities. While this is an effective strategy of accessing gig workers, it may limit the diversity of my sample. Second, ride-hailing and delivery work are male-dominated industries. Additionally, while I found that workers’ economic dependence on the platform might affect their responses to metrics, I did not ask for the interviewees’ platform-based earnings during the

interviews.

Moving forward, future research should collect such data and consider sampling gig workers by demographic characteristics, if possible. Alongside platform-based earnings, workers' education level and age may also affect their reactions to metrics. These two factors may shape the extent to which one considers platform work as a viable future career option. For example, young people with a higher education qualification (e.g., some of the delivery workers I interviewed like Shawn and Harvey) are more likely to consider platform work as a short-term side job rather than a long-term career. In this vein, metrics may not seem important to them. Methodologically, follow-up research can implement a survey before conducting in-depth interviews. The survey may ask respondents to report their attitudes towards labor platforms, evaluation of job quality, previous work experiences, usage of worker-to-worker online communities, and demographics. The survey method should provide a more conformable environment for participants to respond to questions, compared to accepting an interview request from a stranger online. Empirically, the survey can provide an overview of workers' attitudes towards platforms and job quality. It can also allow for recruiting interviews at a later stage of research.

Second, the COVID-19 pandemic has been challenging for gig workers and any research that involves human subjects. For this project, I encountered difficulties in recruiting participants in the early days of the pandemic. I also had to drop the case of platform co-ops and include delivery platforms as the third case. Yet, there are opportunities for thinking about how the pandemic has impacted gig workers' precarious experience, and more broadly, the gig economy. On the one hand, the

pandemic has reportedly put gig workers in a more economically vulnerable position than before. On the other hand, the introduction of contactless delivery and tasks may minimize the interactions between workers and customers, which further allows workers to avoid dealing with difficult customers. At a macro level, the pandemic has increased the demand and supply for delivery services, while it has lowered the demand for ride-hailing services and gig work that require a significant amount of face-to-face interactions. Hence, future research may consider how the pandemic may result in varied precarious experiences in different sectors of the gig economy.

Third, this project focused on digitally-enabled service workers in the United States. While this project offers a comparative assessment of platform-based metrics, it did not examine national and cultural differences. For example, while Uber seems to successfully navigate legal gray zones in favor of their business in the United States, it has difficulties in operating the business in East Asia. In China, Uber entered the market and established an independent business entity in 2014. The company eventually decided to quit the market in 2016 when the Chinese government began legalizing and regulating the ride-hailing industry. In developing countries, governments might decide *not* to effectively regulate labor platforms because they can create employment opportunities (Anwar & Graham, 2021). At the same time, gig workers may earn a higher income, compared to those who work for local companies in developing countries (Anwar & Graham, 2021). As there are up to 55 million gig workers around the world (The Fairwork Project, 2020), future research should account for how national policy regulation shapes the operation of labor platforms as well as workers' practices.

Future Research and Conclusion

Overall, this multi-case study dissertation offers a comparative lens for understanding the role of algorithmic metrics in shaping the trilateral relationship between platform owners, service workers, and customers in the gig economy. Moving forward, comparative research can provide invaluable insights into convergence and divergence across labor platforms, sectors, and countries. First, this project examined the disciplinary impacts of metrics on workers in the contexts of ride-hailing platforms, delivery platforms, and TaskRabbit, but there are other types of platforms in the gig economy. For example, Ticona and Mateescu (2018) argue for the importance of studying carework platforms to counter the "Uberization" narrative in the literature. There are also freelancing platforms for knowledge workers such as Upwork and Fiverr. Hence, future research should explore how the tensions between algorithmic labor control and resistance as well as between the management by and of customers take place across these platforms.

Additionally, though I had difficulties in recruiting participants from platform co-ops, it is still valuable to compare the governance of profit-driven labor platforms and platform co-ops. Scholz (2017) proposes that platform co-ops should avoid incorporating customer ratings to protect workers from excessive surveillance. Instead, platform co-ops should be "co-determined work," meaning that "the design process of labor platforms should involve workers from the first moment and throughout the platform's existence" (p. 181). Gray and Suri (2019) suggest that platform co-ops adopt a "double bottom line" approach to their businesses, hoping to balance profit sharing and co-ownership (p. 147; see also Scholz, 2017; Schneider, 2018). There is

little empirical research on platform co-ops. A notable exception is Schor's (2020) study of Stocksy—a platform co-op governed by photographers. The headquarter of this co-op is based in Canada, but it has members from 65 countries. Essentially, the co-op limits its membership to talented photographers and allows their members to sell their products to businesses. Compared to profit-driven labor platforms, Stocksy has a better compensation structure and its limited membership also avoid intense competition among members. Yet, Schor points out that participation remains a key challenge in the co-op because of time difference, language barriers, and members' lack of interest in taking part in governance. Future research should consider worker-owned co-ops in service sectors such as Up & Go in New York City and Green Taxi in Colorado. As these co-ops avoid using metrics, they can serve as negative cases to consider alternative ways of (e)valuating workers' efforts (Stark, 2009).

Future comparative research can also examine the impacts of metrics, and more broadly, quantification on different sectors. In the context of service work, corporations increasingly rely upon algorithms and metrics to monitor service workers' productivity, automate their work schedule, and determine employment opportunities, particularly in retail and food service industries (Levy & Barocas, 2018). Future research could consider how the introduction of such algorithmic technologies reshape the interplay between store managers, workers, and customers. In this case, store managers and workers may both be subject to the governance of metrics. How do they respond to such metrics and organizational change variously? Under what situations would the interests of managers and workers align with one another? How are algorithmic metrics different from traditional forms of customer

feedback?

Lastly, future research should consider comparing gig workers across countries. To date, the discussion has largely focused on the gig economy in European and North American contexts (e.g., Cant, 2020; Ravenelle, 2019; Rosenblat, 2018a; Schor, 2020; Shapiro, 2018; Ticona & Mateescu, 2018). This is perhaps because the market power is highly concentrated in the hands of a few US-based tech companies.[27] There are a few exceptions. For example, a group of researchers from the University of Oxford has examined algorithmic labor control and resistance in Southeast Asia and Africa (e.g., Anwar & Graham, 2020; Wood et al., 2019). There is also growing body of literature concerning platformization (e.g., Chen & Qiu, 2019; de Kloet et al., 2019) and gig workers (e.g., Chen, 2018; Chen & Sun, 2020; Lei, 2021) in China. Moving forward, I plan to further my research on algorithmic labor control and resistance as well as platform governance in the Chinese contexts. There are at least two potential avenues for conducting comparative research. First, many large labor platforms operate in more than countries (e.g., Uber, DiDi, Deliveroo, etc.). How do they adapt their business models and management of workers to different national and regulatory contexts?[28] This allows us to think further about the varieties of platformization.

The second line of potential future research has to do with how algorithmic

[27] When discussing market concentration of platforms, researchers generally focus on giant corporations including the so-called "GAFAM" (Google, Apple, Facebook, Amazon, and Microsoft) (see, for example, Nieborg & Poell, 2018; van Dijck et al., 2018). Some may also account for the so-called "BAT" (Baidu, Alibaba, and Tencent) in China (see, for instance, Jia & Winseck, 2018). I could not find existing research that specifically maps the global market share of labor platforms, possibly because many of them have not yet been public companies.

[28] Though TikTok (or Douyin in China) is social networking platform rather than labor platform, this is a telling example of how the platform establishes separate entities to adapt to different regulatory contexts.

labor control and resistance take place in different cultural contexts. For instance, how do regulatory contexts affect labor control? How do platforms mobilize cultural values to legitimize their mechanisms of labor control? How do the management by and of customers take place in different cultural contexts? Existing research has shown that there are worker-to-worker online communities in different countries (e.g., Chen, 2018; Gray et al., 2016; Wood et al., 2018). How might we theorize these online communities as sources of individual and collective resistance in the gig economy?

In closing, this dissertation has examined how metrics enable algorithmic labor control and resistance in the gig economy. The affective power of metrics contributes to normalizing and rationalizing customer-oriented rules, which further incites workers to strategically manage their customers. I hope this dissertation provides a first step for me to build a comparative research program on the politics of labor platforms and governance.

REFERENCES

ABC7. (2013, December 19). 7 on your side investigates Lyft's voluntary donations. *ABC7 News*. https://abc7news.com/archive/9366763/

Abril, D. (2020, February 25). Uber Eats boss steps down effective immediately, as the company focuses on profitability. *Fortune*. https://fortune.com/2020/02/25/uber-eats-jason-droege-steps-down/

Ajunwa, I., & Greene, D. (2019). Platforms at work: Automated hiring platforms and other new intermediaries in the organization of work. *Research in the Sociology of Work, 33*, 61–91. https://doi.org/10.1108/S0277-283320190000033005

Ajunwa, I., Crawford, K., & Schultz, J. (2017). Limitless worker surveillance. *California Law Review, 105*(3), 735–770. https://dx.doi.org/10.15779/Z38BR8MF94

Alba, D. (2015, June 22). Instacart shoppers can now choose to be real employees. *Wired*. https://www.wired.com/2015/06/instacart-shoppers-can-now-choose-real-employees/

Anderson, C. (2011). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism, 12*(5), 550–566. https://doi.org/10.1177/1464884911402451

Anwar, M. A., & Graham, M. (2020). Hidden transcripts of the gig economy: Labour agency and the new art of resistance among African gig workers. *Environment and Planning A: Economy and Space, 52*(7), 1269–1291. https://doi.org/10.1177/0308518X19894584

Anwar, M. A., & Graham, M. (2021). Between a rock and a hard place: Freedom, flexibility, precarity and vulnerability in the gig economy in Africa. *Competition & Change, 25*(2), 237–258. https://doi.org/10.1177/1024529420914473

Askay, D. A. (2015). Silence in the crowd: The spiral of silence contributing to the positive bias of opinions in an online review system. *New Media & Society, 17*(11), 1811–1829. https://doi.org/10.1177/1461444814535190

Avedian, S. (2019, December 9). The rating game: Who is better between Uber and Lyft? *The Rideshare Guy*. https://therideshareguy.com/ratings-game/

Bain, P., & Taylor, P. (2000). Entrapped by the 'electronic panopticon'? Worker resistance in the call centre. *New Technology, Work and Employment, 15*(1), 2–18. https://doi.org/10.1111/1468-005X.00061

Ball, K. (2010). Workplace surveillance: An overview. *Labor History, 51*(1), 87–106. https://doi.org/10.1080/00236561003654776

Barley, S. R. (1986). Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Administrative Science Quarterly, 31*(1), 78–108. https://doi.org/10.2307/2392767

Baron, J. N., Dobbin, F. R., & Jennings, P. D. (1986). War and peace: The evolution of modern personnel administration in U.S. industry. *American Journal of Sociology, 92*(2), 350–383. https://doi.org/10.1086/228504

Baym, N. (2018). *Playing to the crowd: Musicians, audiences, and the intimate work of connection*. New York University Press.
Beckert, J. (2013). *Imagined futures: Fictional expectations in the economy. Theory and Society, 42, 219*–240. https://doi.org/10.1007/s11186-013-9191-2
Beer, D. (2016). *Metric power*. Palgrave Macmillan.
Bélanger, J., & Edwards, P. (2013). The nature of front-line service work: Distinctive features and continuity in the employment relationship. *Work, Employment and Society*, *27*(3), 433–450. https://doi.org/10.1177/0950017013481877
Berg, J., & Johnston, H. (2019). Too good to be true? A comment on Hall and Krueger's analysis of labor market for Uber's driver-partners. *ILR Review*, *72*(1), 39–68. https://doi.org/10.1177/0019793918798593
Bhuiyan, J. (2017, August 23). Take a look at Uber's first pitch deck from 2008. *Recode*. https://www.vox.com/2017/8/23/16189048/uber-pitch-deck-2008-ubercab-travis-kalanick-founder-startup
Bieber, F., & Moggia, J. (2020). Risk shifts in the gig economy: The normative case for an insurance scheme against the effects of precarious work. *Journal of Political Philosophy*. Advance online publication. https://doi.org/10.1111/jopp.12233
Bishop, V., & Hoel, H. (2008). The customer is always right? Exploring the concept of customer bullying in the British employment service. *Journal of Consumer Culture*, *8*(3), 341–367. https://doi.org/10.1177/1469540508095303
Blank, G. (2007). *Critics, ratings, and society: The sociology of reviews*. Rowan & Littlefield.
Bolton, S. C., & Houlihan, M. (2005). The (mis)representation of customer service. *Work, Employment and Society*, *19*(4), 685–703. https://doi.org/10.1177/0950017005058054
Bosa, D., & Batchelor, L. (2021, January 7). *Instacart taps Goldman Sachs banker as CFO ahead of expected 2021 IPO. CNBC.* https://www.cnbc.com/2021/01/07/instacart-taps-goldman-sachs-banker-as-cfo-ahead-of-expected-2021-ipo.html
Bowman, C. (2019, November 26). I drive for Uber and Lyft – here's what your driver thinks about you based on your rating. *Business Insider*. https://www.businessinsider.com/uber-lyft-rating-passenger-meaning-2019-11
Braverman, H. (1974). *Labor and monopoly capital: The degradation of work in the twentieth century*. Monthly Review Press.
Brown, E. (2017, April 7). TaskRabbit shops for a buyer. *The Wall Street Journal*. https://www.wsj.com/articles/taskrabbit-shops-for-a-buyer-1491606423?mod=article_inline
Bucher, E. L., Schou, P. K., & Waldkirch, M. (2021). Pacifying the algorithm – Anticipatory compliance in the face of algorithmic management in the gig economy. *Organization*, *28*(1), 44–67. https://doi.org/10.1177/1350508420961531
Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086

Buchheit, P. (2020, December 9). *DoorDash from application to IPO*. https://blog.ycombinator.com/doordash-from-application-to-ipo/
Burawoy, M. (1979). *Manufacturing consent: Changes in the labor process under monopoly capitalism*. University of Chicago Press.
Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951715622512
Busque, L. (2012, July 31). The role (and future) of social reputation. *HuffPost*. https://www.huffpost.com/entry/online-shopping-trust_b_1556885
Buzzard, K. (2012). *Tracking the audience: The ratings industry from analog to digital*. Routledge.
Callaghan, G., & Thompson, P. (2001). Edwards revisited: Technical control and call centres. *Economic and Industrial Democracy*, *22*(13), 13–37. https://doi.org/10.1177/0143831X01221002
Campbell, H. (2015, August 31). My rating system for Uber and Lyft passengers. *The Rideshare Guy*. https://therideshareguy.com/my-rating-system-for-uber-and-lyft-passengers/
Campbell, H. (2018). *The rideshare guide: Everything you need to know about driving for Uber, Lyft, and other ridesharing companies*. Skyhorse Publishing.
Campbell, H. (2021, February 24). Lyft & Uber driver survey 2019: Uber driver satisfaction takes a big hit. *The Rideshare Guy*. https://therideshareguy.com/uber-driver-survey/
Cant, C. (2020). *Riding for Deliveroo: Resistance in the new economy*. Polity.
Captain, S. (2019, July 16). Instacart drivers fear that hiding customer address puts them at risk. *Fast Company*. https://www.fastcompany.com/90377506/instacart-drivers-worry-that-new-policy-hiding-some-customer-addresses-may-put-them-at-risk
Carruthers, B. G. (2013). From uncertainty toward risk: The case of credit ratings. *Socio-Economic Review*, *11*(3), 525–551. https://doi.org/10.1093/ser/mws027
Chen, J. Y. (2018). Thrown under the bus and outrunning it! The logic of Didi and taxi drivers' labour and activism in the on-demand economy. *New Media & Society*, *20*(8), 2691–2711. https://doi.org/10.1177/1461444817729149
Chen, J. Y., & Qiu, J. L. (2019). Digital utility: Datafication, regulation, labor, and DiDi's platformization of urban transport in China. *Chinese Journal of Communication*, *12*(3), 274–289. https://doi.org/10.1080/17544750.2019.1614964
Chen, J. Y., & Sun, P. (2020). Temporal arbitrage, fragmented rush, and opportunistic behaviors: The labor politics of time in the platform economy. *New Media & Society*, *22*(9), 1561–1579. https://doi.org/10.1177/1461444820913567
Cheng, M. (2019, October 17). *How universities are starting to prepare students for the gig economy*. *Quartz*. https://qz.com/work/1724712/how-universities-are-preparing-students-for-the-gig-economy/
Christin, A. (2017). Algorithms in practice: Comparing web journalism and criminal justice. *Big Data & Society*, *4*(2). https://doi.org/10.1177/2053951717718855

Christin, A. (2018). Counting clicks: Quantification and variation in web journalism in the United States and France. *American Journal of Sociology*, *123*(5), 1382–1415. https://doi.org/10.1086/696137

Christin, A. (2020a). *Metrics at work: Journalism and the contested meaning of algorithms.* Princeton University Press.

Christin, A. (2020b). Algorithmic ethnography, during and after COVID-19. *Communication and the Public*, *5*(3-4), 108–111. https://doi.org/10.1177/2057047320959850

Christin, A., & Petre, C. (2020). Making peace with metrics: Relational work in online news production. *Sociologica, 14(2),* 133–156. https://doi.org/10.6092/issn.1971-8853/11178

Creswell, J.W. (2014). *Research design: Qualitative, quantitative, and mixed method approaches* (4th Ed.) Sage.

Culpepper, P. D., & Thelen, K. (2020). Are we all Amazon primed? Consumers and the politics of platform power. *Comparative Political Studies*, *53*(2), 288–318. https://doi.org/10.1177/0010414019852687

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2020). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly*, *65*(3), 644–676. https://doi.org/10.1177/0001839219867024

Cutler, K.M. (2012, May 22). Zimride's Lyft is going to give Uber some lower-priced competition. *TechCrunch*. https://techcrunch.com/2012/05/22/zimrides-lyft-is-going-to-give-uber-some-lower-priced-competition/

Damiani, J. (2020, June 9). *Instacart surges past Walmart in online grocery market. Forbes.* https://www.forbes.com/sites/jessedamiani/2020/06/09/instacart-surges-past-walmart-in-online-grocery-market/

Davis, F. (1959). The cabdriver and his fare: Facets of a fleeting relationship. *American Journal of Sociology*, *65*(2), 158–165. https://doi.org/10.1086/222656

de Kloet, J., Poell, T., Zeng, G., & Chow, Y. F. (2019). The platformization of Chinese Society: infrastructure, governance, and practice. *Chinese Journal of Communication*, *12*(3), 249–256. https://doi.org/10.1080/17544750.2019.1644008

De Stefano, V. (2016). *The rise of the "just-in-time workforce": On-demand work, crowdwork and labour protection in the "gig-economy."* International Labour Office.

Del Rey, J. (2017, February 20). Instacart is playing games with its workers' pay — and will eventually suffer for it. *Vox.* https://www.vox.com/2017/2/20/14503128/instacart-service-fee-tips-controversy

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms *Management Science*, *49*(10), 1275–1444. https://doi.org/10.1287/mnsc.49.10.1407.17308

Dellarocas, C., & Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science*, *54*(3), 460–476. https://doi.org/10.1287/mnsc.1070.0747

Deng, Y. (2016, February 9). How Instacart's pricing changes impact retailers. *Supermarket News*. https://www.supermarketnews.com/blog/how-instacarts-pricing-changes-impact-retailers

Denzin, N. K., & Lincoln, Y. S. (2011). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (Fourth ed., pp. 1–20). Sage.

DoorDash. (n.d.-a). *Become a driver & deliver with DoorDash*. https://www.doordash.com/dasher/signup/

DoorDash (n.d.-b). *Deactivation policy - United States*. https://help.doordash.com/dashers/s/deactivation-policy-us?language=en_US

DoorDash. (2017a, April 10). *An update on Dasher class action*. Medium. https://medium.com/@DoorDash/an-update-on-dasher-class-action-6990643be500

DoorDash. (2017b, June 21). *Four years in and just getting started.* Medium. https://blog.doordash.com/four-years-in-and-just-getting-started-9a8ab3583c37

DoorDash. (2019, November 12). *Update on new pay model: More dollars to Dashers.* Medium. https://blog.doordash.com/update-on-new-pay-model-more-dollars-to-dashers-28454459d9a2

DoorDash, Inc. [DoorDash]. (2020a). Form S-1 registration statement. https://www.sec.gov/Archives/edgar/data/1792789/000119312520292381/d752207ds1.htm

DoorDash. (2020b). *Making ratings more actionable*. https://help.doordash.com/dashers/s/article/Making-Ratings-More-Actionable?language=en_US

Duffy, B. E. (2017). *(Not) getting paid to do what you love: Gender, social media, and aspirational work.* Yale University Press.

Duffy, B. E. (2020). Algorithmic precarity in cultural work. *Communication and the Public*, *5*(3-4), 103–107. https://doi.org/10.1177/2057047320959855

Dzieza, J. (2015, October 28). *The rating game: How Uber and its peers turned us into horrible bosses. The Verge*. https://www.theverge.com/2015/10/28/9625968/rating-system-on-demand-economy-uber-olive-garden

Edwards, R. (1984). Forms of control in the labor process: A historical analysis. In F. Fischer & C. Sirianni (Eds.), *Critical studies in organization and bureaucracy* (pp. 109–142). Temple University Press.

Eisenhardt, K. M. (1989). Building theories from case study research. *The Academy of Management Review*, *14*(4), 532–550. https://doi.org/10.2307/258557

Ellway, B. P. W. (2013). Making it personal in a call centre: Electronic peer surveillance. *New Technology, Work and Employment*, *28*(1), 37–50. https://doi.org/10.1111/ntwe.12002

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, *113*(1), 1–40. https://doi.org/10.1086/517897

Espeland, W. N., & Sauder, M. (2016). *Engines of anxiety: Academic rankings, reputation, and accountability*. Russell Sage Foundation.

Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a social process. *Annual Review of Sociology*, *24*, 313–343. https://doi.org/10.1146/annurev.soc.24.1.313

Espeland, W. N., & Stevens, M. L. (2008). A sociology of quantification. *European Journal of Sociology*, *49*(3), 401–436. https://doi.org/10.1017/S0003975609000150

Esposito, E., & Stark, D. (2019). What's observed in a rating? Rankings as orientation in the face of uncertainty. *Theory, Culture & Society*, *36*(4), 3–26. https://doi.org/10.1177/0263276419826276

Ettlinger, N. (2018). Algorithmic affordances for productive resistance. *Big Data & Society*, *5*(1). https://doi.org/10.1177/2053951718771399

Farrell, D., Greig, F., & Hamoudi, A. (2018). *The online platform economy in 2018: Drivers, workers, sellers, and lessors*. JP Morgan Chase & Co. Institute. https://www.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/institute/pdf/institute-ope-2018-exec-summary.pdf

Feiner, L. (2020, November 13). *DoorDash releases filing to go public, reports $149 million in losses on revenue of $1.9 billion through September*. *CNBC*. https://www.cnbc.com/2020/11/13/doordash-releases-s-1-for-ipo.html

Ferrari, F., & Graham, M. (2021). Fissures in algorithmic power: Platforms, code, and contestation. *Cultural Studies*. Advance online publication. https://doi.org/10.1080/09502386.2021.1895250

Figart, D. M. (2001). Wage-setting under Fordism: The rise of job evaluation and the ideology of equal pay. *Review of Political Economy*, *13*(4), 405–425. https://doi.org/10.1080/0953825012009935

Fleming, F. (2017). The human capital hoax: Work, debt and insecurity in the ear of Uberization. *Organization Studies*, *38*(5), 691–709. https://doi.org/10.1177/0170840616686129

Fleming, P., & Sturdy, A. (2011). 'Being yourself' in the electronic sweatshop: New forms of normative control. *Human Relations*, *64*(2), 177–200. https://doi.org/10.1177/0018726710375481

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*(2), 219–245. https://doi.org/10.1177/1077800405284363

Fortner, R. S., & Christians, C. G. (1989). Separating wheat from chaff in qualitative studies. In G. H. Stempel III & B. H. Westley (Eds.), *Research methods in mass communication* (Second ed., pp. 375–387). Prentice Hall.

Fourcade, M. (2011). Cents and sensibility: Economic valuation and the nature of "nature." *American Journal of Sociology*, *116*(6), 1721–1777. https://doi.org/10.1086/659640

Fourcade, M., & Healy, K. (2017). Seeing like a market. *Socio-Economic Review*, *15*(1), 9–29. https://doi.org/10.1093/ser/mww033

Frank, B. H. (2013, June 17). Q&A: Instacart CEO on competing with Amazon Fresh and the future of grocery shopping. *GeekWire*. https://www.geekwire.com/2013/qa-instacart-ceo-apoorva/

Frenkel, S., Korczynski, M., Donoghue, L., & Shire, K. (1995). Re-constituting work: Trends towards knowledge work and info-normative control. *Work,*

Employment and Society, *9*(4), 773–796. https://doi.org/10.1177/095001709594008

Freytas-Tamura, K. de. (2020, December 9). Food delivery apps are booming. Their workers are often struggling. *The New York Times*. https://www.nytimes.com/2020/11/30/nyregion/bike-delivery-workers-covid-pandemic.html

Fuller, L., & Smith, V. (1991). Consumers' reports: Management by customers in a changing economy. *Work, Employment and Society*, *5*(1), 1–16. https://doi.org/10.1177/0950017091005001002

Gallagher, B. (2013, July 26). Y Combinator-backed DoorDash delivers food quickly in South Bay, hopes to expand beyond food. *TechCrunch*. https://techcrunch.com/2013/07/26/y-combinator-backed-doordash-delivers-food-quickly-in-south-bay-hopes-to-expand-beyond-food/

Gandini, A. (2016). Digital work: Self-branding and social capital in the freelance knowledge economy. *Marketing Theory*, *16*(1), 123–141. https://doi.org/10.1177/1470593115607942

Gandini, A. (2019). Labour process theory and the gig economy. *Human Relations*, *72*(6), 1039–1056. https://doi.org/10.1177/0018726718790002

Gebel, M. (2020, November 18). I'm a DoorDash delivery driver who makes anywhere from $200 to $800 a week. Here's what I wish I knew before signing up to drive for the company. *Business Insider*. https://www.businessinsider.com/doordash-delivery-driver-in-oregon-what-i-wish-i-knew-2020-11

Gerlitz, C., & Lury, C. (2014). Social media and self-evaluating assemblages: On numbers, orderings and values. *Distinktion: Journal of Social Theory*, *15*(2), 174–188. https://doi.org/10.1080/1600910X.2014.920267

Gig Economy Data Hub. (2020). *How many gig workers are there?* https://www.gigeconomydata.org/basics/how-many-gig-workers-are-there

Gillespie, T. (2010). The politics of 'platforms.' *New Media & Society*, *12*(3), 347–364. https://doi.org/10.1177/1461444809342738

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–194). MIT Press.

Glaser, A. (2019, November 5). Instacart workers are striking because of the app's user interface. *Slate*. https://slate.com/business/2019/11/instacart-workers-striking-app-user-interface-dark-pattern-design.html

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine Publishing Company.

Goldin, R. (2016, November 18). *A beginner's guide to Uber Everything*. Medium. https://medium.com/uber-design/a-beginners-guide-to-uber-everything-f981392a43df

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, *22*(6), 854–871. https://doi.org/10.1080/1369118X.2019.1573914

Gottfried, H. (1991). Mechanisms of control in the temporary help service industry. *Sociological Forum*, *6*(4), 699–713.

Gray, M. L., & Suri, S. (2019). *Ghost work: How to stop Silicon Valley from building a new global underclass*. Houghton Mifflin Harcourt.

Gray, M. L., Suri, S., Ali, S. S., & Kulkarni, D. (2016). The crowd is a collaborative network. *CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 134–147. https://doi.org/10.1145/2818048.2819942

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274. https://doi.org/10.3102/01623737011003255

Gregory, K. (2021). 'My life is more valuable than this': Understanding risk among on-demand food couriers in Edinburgh. *Work, Employment and Society*, *35*(2), 316–331. https://doi.org/10.1177/0950017020969593

Griesbach, K., Reich, A., Elliott-Negri, L., & Milkman, R. (2019). Algorithmic control in platform food delivery work. *Socius*, *5*, 1–15. https://doi.org/10.1177/2378023119870041

Griswold, A. (2016, April 18). Uber is killing off its instant food delivery option in New York City. *Quartz*. https://qz.com/664552/uber-is-killing-off-its-instant-food-delivery-option-in-new-york-city/

Gurley, L. K. (2019, November 8). Instacart cuts 'quality' bonus after workers go on 3-day strike. *Vice*. https://www.vice.com/en/article/j5y3q7/instacart-cuts-quality-bonus-after-workers-go-on-3-day-strike

Gurley, L. K. (2020, February 1). Instacart workers win historic union election. *Vice*. https://www.vice.com/en/article/qjdvgm/instacart-workers-win-historic-union-election

Hall, J. V., & Krueger, A. B. (2018). An analysis of the labor market for Uber's driver-partners in the United States. *ILR Review*, *71*(3), 705–732. https://doi.org/10.1177/0019793917717222

Hall, R. (2010). Renewing and revising the engagement between labour process theory and technology. In P. Thompson & C. Smith (Eds.), *Working life: Renewing labour process analysis* (pp. 159–181). Palgrave Macmillan.

Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M., & Wilson, C. (2017). Bias in online freelance marketplaces: Evidence from TaskRabbit and Fiverr. *CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1914–1933. https://doi.org/10.1145/2998181.2998327

Hartanto, Y., & Attwell, B. (2019, May 29). Activity/service as a dependency: Rethinking Android architecture for the Uber driver app. *Uber Engineering*. https://eng.uber.com/activity-service-dependency-android-app-architecture/

Hawkins, A. J. (2018, April 10). Uber redesigned its driver app with input from actual drivers. *The Verge*. https://www.theverge.com/2018/4/10/17215298/uber-redesign-driver-app-promotions-gamify

Hawkins, A. J. (2019a, March 12). Uber settles driver classification lawsuit for $20 million. *The Verge*. https://www.theverge.com/2019/3/12/18261755/uber-driver-classification-lawsuit-settlement-20-million

Hawkins, A. J. (2019b, September 26). Uber overhauls its app in ambitious bid to become 'the operating system for your everyday life.' *The Verge*. https://www.theverge.com/2019/9/26/20884800/uber-app-announcement-ubereats-safety-transit-updates

Hawkins, A. J. (2020a, August 12). *Lyft joins Uber in threatening to pull out of California over driver status. The Verge*. https://www.theverge.com/2020/8/12/21365518/lyft-threatens-shut-down-california-q2-2020

Hawkins, A. J. (2020b, November 4). Uber and Lyft had an edge in the Prop 22 fight: Their apps. *The Verge*. https://www.theverge.com/2020/11/4/21549760/uber-lyft-prop-22-win-vote-app-message-notifications

Hearn, A. (2010). Structuring feeling: Web 2.0, online ranking and rating, and the digital 'reputation' economy. *Ephemera, 10*(3/4), 421–438.

Hochschild, A. (1983/2012). *The managed heart: Commercialization of human feeling*. University of California Press.

Hodson, R. (1995). Worker resistance: An underdeveloped concept in the sociology of work. *Economic and Industrial Democracy, 16*(1), 79–110. https://doi.org/10.1177/0143831X9501600104

Horowitz, S. (2011, September 1). *The freelance surge is the industrial revolution of our time. The Atlantic.* https://www.theatlantic.com/business/archive/2011/09/the-freelance-surge-is-the-industrial-revolution-of-our-time/244229/

Hoskin, K. (1996). The "awful idea of accountability": Inscribing people into the measurement of objects. In R. Munro & J. Mouritsen (Eds.), *Accountability: Power, ethos and the technologies of managing* (pp. 265–282). International Thomson Business Press.

How to Start a Startup. (2014, October 16). *Lecture 8 - How to get started, doing things that don't scale, Press* [Video]. YouTube. https://www.youtube.com/watch?v=oQOC-qy-GDY&ab_channel=HowtoStartaStartup

Howcroft, D., & Bergvall-Kåreborn, B. (2019). A typology of crowdwork platforms. *Work, Employment and Society, 33*(1), 21–38. https://doi.org/10.1177/0950017018760136

Humphreys, L. (2020). Qualitative sampling and Internet research. In E. Hargittai (Ed.), *Research exposed: How empirical social science gets done in the digital age* (pp. 78–100). Columbia University Press.

Hyman, R. (1987). Strategy or structure? Capital, labour and control. *Work, Employment and Society, 1*(1), 25–55. https://doi.org/10.1177/0950017087001001004

Ikeler, P. (2016). *Hard sell: Work and resistance in retail chain*s. Cornell University Press.

Ince, J. (2017, March 29). How to fix Uber and Lyft's rating system. *The Rideshare Guy*. https://therideshareguy.com/how-to-fix-uber-and-lyfts-rating-system/

Instacart. (2016, September 22). *We're making updates to improve our shopper experience*. Medium. https://news.instacart.com/were-making-updates-to-improve-our-shopper-experience-200f4b680d07

Instacart. (2018, December 13). *Updates to our partnership with Whole Foods Market*. Medium. https://medium.com/shopper-news/update-to-our-partnership-with-whole-foods-market-f114eda737f

Instacart. (2020a, March 16). *New updates and support for the shopper community*. Medium. https://medium.com/shopper-news/new-updates-and-support-for-the-shopper-community-ad757ffdecd5

Instacart. (2020b, June 9). *Maximizing your access to batches*. Medium. https://medium.com/shopper-news/maximizing-your-access-to-batches-d38a9e943eaa

Instacart. (2020c, November 23). *Introduction to ratings | Getting started | Instacart* [Video]. YouTube. https://www.youtube.com/watch?v=ogknAHBPVsE&ab_channel=Instacart

Irani, L. (2015). The cultural work of micro-work. *New Media & Society, 17*(5), 720–739. https://doi.org/10.1177/1461444813511926

Issac, M. (2019). *Super pumped: The battle for Uber*. W. W. Norton & Company.

Jia, L., & Winseck, D. (2018). The political economy of Chinese internet companies: Financialization, concentration, and capitalization. *International Communication Gazette*, *80*(1), 30–59. https://doi.org/10.1177/1748048517742783

Kalleberg, A. L. (2009). Precarious work, insecure workers: Employment relations in transition. *American Sociological Review*, *74*(1), 1–22. https://doi.org/10.1177/000312240907400101

Kalleberg, A. L., & Dunn, M. (2016). Good jobs, bad jobs in the gig economy. *Perspectives on Work, 20,* 10–13, 74–75.

Kalleberg, A. L., & Vallas, S. P. (2018). Probing precarious work: Theory, research, and politics. In A. L. Kalleberg & S. P. Vallas (Eds.), *Precarious work: Causes, characteristics, and consequences* (pp. 1–30). Emerald.

Karls, N. (2019, November 14). Instacart cut bonuses after a strike. Workers suspect it's retaliation. *Salon*. https://www.salon.com/2019/11/13/instacart-cut-bonuses-after-a-strike-workers-suspect-its-retaliation/

Kässi, O., & Lehdonvirta, V. (2018). Online labour index: Measuring the online gig economy for policy and research. *Technological Forecasting and Social Change*, *137*, 241–248. https://doi.org/10.1016/j.techfore.2018.07.056

Kear, M. (2017). Playing the credit score game: Algorithms, 'positive' data and the personification of financial objects. *Economy and Society*, *46*(3–4), 346–368. https://doi.org/10.1080/03085147.2017.1412642

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Kirsner, S. (2013, December 11). Instacart introduces neighbor-to-neighbor grocery delivery to Boston. *Boston.com*.

http://archive.boston.com/business/technology/innoeco/2013/12/instacart_introduces_neighbor-.html
Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information Communication and Society*, *20*(1), 14–29. https://doi.org/10.1080/1369118X.2016.1154087
Klawitter, E. F. (2020). Stitching data: A multimodal approach to learning about independent artists' social media use. In E. Hargittai (Ed.), *Research exposed: How empirical social science gets done in the digital age* (pp. 205–224). Columbia University Press.
Korczynski, M. (2004). Back-office service work: Bureaucracy challenged? *Work, Employment and Society*, *18*(1), 97–114. https://doi.org/10.1177/0950017004040764
Korczynski, M. (2007). Service work, social theory, and collectivism: A reply to Brook. *Work, Employment and Society*, *21*(3), 577–588. https://doi.org/10.1177/0950017007080019
Korczynski, M. (2009a). The mystery customer: Continuing absences in the sociology of service work. *Sociology*, *43*(5), 952–967. https://doi.org/10.1177/0038038509340725
Korczynski, M. (2009b). Understanding the contradictory lived experience of service work: The customer-oriented bureaucracy. In M. Korczynski & C. L. Macdonald (Eds.), *Service work: Critical perspectives* (pp. 73–91). Routledge.
Korczynski, M. (2013). The customer in the sociology of work: Different ways of going beyond the management-worker dyad. *Work, Employment and Society*, *27*(6), NP1–NP7. https://doi.org/10.1177/0950017012464424
Kuehn, K. M. (2013). "There's got to be a review democracy": communicative capitalism, neoliberal citizenship and the politics of participation on the consumer evaluation website Yelp.com. *International Journal of Communication*, *7*, 607–625.
Kvale, S. (2002). The social construction of validity. In N. K. Denzin & Y. S. Lincoln (Eds.), *The qualitative inquiry reader* (pp. 299–325). Sage.
Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology*, *38*, 201–221. https://dx.doi.org/10.1146/annurev-soc-070308-120022
Lamont, M., & Swidler, A. (2014). Methodological pluralism and the possibilities and limits of interviewing. *Qualitative Sociology*, *37*(2), 153–171. https://doi.org/10.1007/s11133-014-9274-z
Lampe, C. (2012). The role of reputation systems in managing online communities. In H. Masum & M. Tovey (Eds.), *The reputation society: How online opinions are reshaping the offline world* (pp. 77–88). MIT Press.
LaRossa, R. (2005). Grounded theory methods and qualitative family research. *Journal of Marriage and Family*, *67*(4), 837–857. https://doi.org/10.1111/j.1741-3737.2005.00179.x
Lavietes, M., & McCoy, M. (2020, October 19). Waiting for work: pandemic leaves U.S. gig workers clamoring for jobs. *Reuters.*

https://www.reuters.com/article/us-biggerpicture-health-coronavirus-gigw-idUSKBN2741DM
Lee, D. (2018, August 8). Uber Eats is changing its flat fees to delivery fees based on distance. *The Verge*. https://www.theverge.com/2018/8/8/17664742/uber-eats-delivery-fee-distance-flat-booking-fee
Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. In B. Begole & J. Kim (Eds.), *CHI '15 proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 1603–1612). ACM.
Lehdonvirta, V. (2018). Flexibility in the gig economy: Managing time on three online piecework platforms. *New Technology, Work and Employment*, *33*(1), 13–29. https://doi.org/10.1111/ntwe.12102
Lei, Y.W. (2021). Delivering solidarity: Platform architecture and collective contention in China's platform economy. *American Sociological Review*. Advance online publication. https://doi.org/10.1177/0003122420979980
Leidner, R. (1993). *Fast food, fast talk: Service work and the routinization of everyday life*. University of California Press.
Levin, S. (2017, June 27). Uber's scandals, blunders and PR disasters: the full list. *The Guardian*. https://www.theguardian.com/technology/2017/jun/18/uber-travis-kalanick-scandal-pr-disaster-timeline
Levy, A. (2020, December 12). *Three of the 10 biggest tech IPOs have happened in 2020—Including two in the last week. CNBC*. https://www.cnbc.com/2020/12/12/10-biggest-tech-ipos-ever-three-in-2020-two-last-week.html
Levy, K. E. C. (2015). The contexts of control: Information, power, and truck-driving work. *Information Society*, *31*(2), 160–174. http://dx.doi.org/10.1080/01972243.2015.998105
Levy, K. E. C., & Barocas, C. (2018). Refractive surveillance: Monitoring customers to manage workers. *International Journal of Communication*, *12*, 1166–1188.
Levy, S. (2015, November 9). DoorDash wants to own the last mile. *Wired*. https://www.wired.com/2015/11/doordash-wants-to-own-the-last-mile/
Lieber, C. (2019, May 14). Grocery startup Instacart says it's fixed its payment problems. This gig worker says it hasn't. *Vox*. https://www.vox.com/the-goods/2019/5/14/18566237/instacart-shopper-tip-grocery-delivery-payment
Lopez, S. H. (2010). Workers, managers, and customers: Triangles of power in work communities. *Work and Occupations*, *37*(3), 251–271. https://doi.org/10.1177/0730888410375683
Luce, I. D. (2020, December 11). DoorDash's founders share the food delivery company's origin story. *The Business of Business*. https://www.businessofbusiness.com/articles/Doordash-ipo-stock-founders-tony-xu/
Lyft. (n.d.). *Acceptance rate*. https://help.lyft.com/hc/en-us/articles/115013077708
Lyft. (2018, November 15). *For drivers from drivers: New features and improvements coming every month*. https://www.lyft.com/blog/posts/for-drivers-from-drivers-new-features-and-improvements-coming-every-month

Lyft, Inc. [Lyft]. (2019). Form S-1 registration statement. https://www.sec.gov/Archives/edgar/data/1759509/000119312519059849/d633517ds1.htm

Madrigal, A. C. (2019, March 6). *The servant economy*. *The Atlantic*. https://www.theatlantic.com/technology/archive/2019/03/what-happened-uber-x-companies/584236/

Maffie, M. D. (2020a). The role of digital communities in organizing gig workers. *Industrial Relations: A Journal of Economy and Society*, *59*(1), 123–149. https://doi.org/10.1111/irel.12251

Maffie, M. D. (2020b). The perils of laundering control through customers: A study of control and resistance in the ride-hailing industry. *ILR Review*. Advance online publication. https://doi.org/10.1177/0019793920972679

Mahdawi. (2016, March 15). *Are you ready for a future where we're all reviewed like Uber drivers? The Guardian*. http://www.theguardian.com/technology/commentisfree/2016/mar/15/rating-culture-score-personal-privacy-uber-sharing-economy

Maiers, C. (2017). Analytics in action: Users and predictive data in the neonatal intensive care unit. *Information, Communication & Society*, *20*(6), 915–929. https://doi.org/10.1080/1369118X.2017.1291701

Malin, B. J., & Chandler, C. (2017). Free to work anxiously: Splintering precarity among drivers for Uber and Lyft. *Communication, Culture & Critique*, *10*(2), 382–400. https://doi.org/10.1111/cccr.12157

Marshall, A. (2020, March 30). The Covid-19 pandemic aggravates disputes around gig work. *Wired*. https://www.wired.com/story/covid-19-pandemic-aggravates-disputes-gig-work/

Marshall, A., & Baker, G. (2020, March 11). Coronavirus exposes workers to the risks of the gig economy. *Wired*. https://www.wired.com/story/coronavirus-exposes-workers-risks-gig-economy/

Mascali, N. M. (2017, July 12). UberEATS goes 24/7 in New York City. *Metro*. https://www.metro.us/ubereats-goes-24-7-in-new-york-city/

Mascarenhas, N. (2020, March 23). Instacart will bring on 300,000 new full-service shoppers due to coronavirus. *TechCrunch*. https://social.techcrunch.com/2020/03/23/instacart-will-bring-on-300000-new-full-time-shoppers-due-to-coronavirus/

Masum, H., & Tovey, M. (Eds.) (2011). *The reputation society: How online opinions are reshaping the offline world*. MIT Press.

Matsakis, L. (2016, May 24). This guy quit his engineer job to run a blog all about Uber: Harry Campbell is the Rideshare Guy. *Motherboard*. https://motherboard.vice.com/en_us/ article/8q88d5/harry-campbell-the-rideshare-guy

Mau, S. (2019). *The metric society: On the quantification of the social*. Polity.

McDonald, P., Williams, P., & Mayes, R. (2020). Means of control in the organization of digitally intermediated care work. *Work, Employment and Society*. Advance online publication. https://doi.org/10.1177/0950017020969107

McGovern, M. (2017). *Thriving in the gig economy: How to capitalize and compete in the new world of work*. Career Press.
Mears, A. (2015). Working for free in the VIP: Relational work and the production of consent. *American Sociological Review*, *80*(6), 1099–1122. https://doi.org/10.1177/0003122415609730
Mellet, K., Beauvisage, T., Beuscart, J.-S., & Trespeuch, M. (2014). A "democratization" of markets? Online consumer reviews in the restaurant industry. *Valuation Studies*, *2*(1), 125–146. https://doi.org/10.3384/vs.2001-5992.14215
Mennicken, A., & Espeland, W. N. (2019). What's new with numbers? Sociological approaches to the study of quantification. *Annual Review of Sociology*, *45*(1), 223–245. https://doi.org/10.1146/annurev-soc-073117-041343
Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research*, *68*(2), 379–396.
Molla, R. (2019, March 6). *Why companies like Lyft and Uber are going public without having profits. Vox*. https://www.vox.com/2019/3/6/18249997/lyft-uber-ipo-public-profit
Moore, P. V. (2018). *The quantified self in precarity: Work, technology and what counts*. Routledge.
Moore, P. V., & Joyce, S. (2020). Black box or hidden abode? The expansion and exposure of platform work managerialism. *Review of International Political Economy*, *27*(4), 926–948. https://doi.org/10.1080/09692290.2019.1627569
Mosseri, S. (2020). Being watched and being seen: Negotiating visibility in the NYC ride-hail circuit. *New Media & Society*. Advance online publication. https://doi.org/10.1177/1461444820966752
Mulcahy, D. (2016). *The gig economy: The complete guide to getting better work, taking more time off, and financing the life you want*. AMACOM.
Mulcahy, D. (2019). *Universities should be preparing students for the gig economy. Harvard Business Review*. https://hbr.org/2019/10/universities-should-be-preparing-students-for-the-gig-economy
Mulholland, K. (2004). Workplace resistance in an Irish call centre: Slammin', scammin' smokin' an' leavin'. *Work, Employment and Society*, *18*(4), 709–724. https://doi.org/10.1177/0950017004048691
Muller, J. Z. (2018). *The tyranny of metrics*. Princeton University Press.
Narula, S. K. (2015, January 28). The aggressive delivery startup that's proving it pays to play well with others. *Quartz*. https://qz.com/323310/the-aggressive-delivery-startup-thats-proving-it-pays-to-play-well-with-others/
Neilson, B., & Rossiter, N. (2008). Precarity as a political concept, or, Fordism as exception. *Theory, Culture & Society*, *25*(7-8), 51–72. https://doi.org/10.1177/0263276408097796
Newton, C. (2013, May 23). *Temping fate: Can TaskRabbit go from side gigs to real jobs? The Verge*. https://www.theverge.com/2013/5/23/4352116/taskrabbit-temp-agency-gig-economy
Newton, C. (2014, June 17). TaskRabbit is blowing up its business model and becoming the Uber for everything. *The Verge*.

https://www.theverge.com/2014/6/17/5816254/taskrabbit-blows-up-its-auction-house-to-offer-services-on-demand
Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society, 20*(11), 4275–4292. https://doi.org/10.1177/1461444818769694
O'Brien, S. A. (2020, July 30). *Some Instacart shoppers are having their jobs cut during the pandemic*. *CNN*. https://www.cnn.com/2020/07/30/tech/instacart-aldi-sprouts-in-store-shoppers/index.html
O'Connell, J. (1993). Metrology: The creation of universality by the circulation of particulars. *Social Studies of Science, 23*(1), 129–173. https://doi.org/10.1177/030631293023001005
Olanoff, D. (2012, September 17). Lyft's focus on community and the story behind the pink mustache. *TechCrunch*. https://techcrunch.com/2012/09/17/lyfts-focus-on-community-and-the-story-behind-the-pink-mustache/
Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
Petre, C., Duffy, B. E., & Hund, E. (2019). "Gaming the system": Platform paternalism and the politics of algorithmic visibility. *Social Media + Society, 5*(4). https://doi.org/10.1177/2056305119879995
Pollman, E., & Barry, J. M. (2017). Regulatory entrepreneurship. *Southern California Law Review, 90*(3), 383–448.
Pollock, N., D'Adderio, L., Williams, R., & Leforestier, L. (2018). Conforming or transforming? How organizations respond to multiple rankings. *Accounting, Organizations and Society, 61*, 55–68. https://doi.org/10.1016/j.aos.2017.11.003
Porter, T. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.
Prassl, J. (2018). *Humans as a service: The promise and perils of work in the gig economy*. Oxford University Press.
Price, R. (2016). Controlling routine front line service workers: An Australian retail supermarket case. *Work, Employment and Society, 30*(6), 915–931. https://doi.org/10.1177/0950017015601778
Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 469–481. https://doi.org/10.1145/3351095.3372828
Ragin, C. (1992). Introduction: Cases of "what is a case?" In C. Ragin & H. Becker (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 1–18). Cambridge University Press.
Ragin, C. (2009). Reflections on casing and case-oriented research. In D. Byrne & C. Ragin (Eds.), *The Sage handbook of case-based methods* (pp. 522–534). Sage.
Rahman, K. S., & Thelen, K. (2019). The rise of the platform business model and the transformation of twenty-first-century capitalism. *Politics & Society, 47*(2), 177–204. https://doi.org/10.1177/0032329219838932

Raval, N., & Dourish, P. (2016). Standing out from the crowd: Emotional labor, body labor, and temporal labor in ridesharing. In D. Gergle & M. R. Morris (Eds.), *CSCW '16 proceedings of the 19th ACM conference on computer-supported cooperative work & social computing* (pp. 97–107). ACM.

Ravenelle, A. J. (2019). *Hustle and gig: Struggling and surviving in the sharing economy*. University of California Press.

Reid-Musson, E., MacEachen, E., & Bartel, E. (2020). 'Don't take a poo!': Worker misbehaviour in on-demand ride-hail carpooling. *New Technology, Work and Employment, 35*(2), 145–161. https://doi.org/10.1111/ntwe.12159

Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029. https://doi.org/10.1162/154247603322493212

Rosenblat, A. (2018a). *Uberland: How algorithms are rewriting the rules of work*. University of California Press.

Rosenblat, A. (2018b, January 9). The network Uber drivers built. *Fast Company*. https://www.fastcompany.com/40501439/the-network-uber-drivers-built

Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication*, *10*, 3758–3784.

Rosenblat, A., Levy, K., Barocas, S., & Hwang, T. (2017). Discriminating tastes: Uber's customer ratings as vehicles for workplace discrimination. *Policy & Internet*, *9*, 256–279. https://doi.org/10.1002/poi3.153

Roy, K., Zvonkovic, A., Goldberg, A., Sharp, E., & LaRossa, R. (2015). Sampling richness and qualitative integrity: Challenges for research with families. *Journal of Marriage and Family*, *77*(1), 243–260. https://doi.org/10.1111/jomf.12147

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854–856. https://doi.org/10.1126/science.1121066

Sallaz, J. J. (2002). The house rules: Autonomy and interest among service workers in the contemporary casino industry. *Work and Occupations*, *29*(4), 394–427. https://doi.org/10.1177/0730888402029004002

Sauder, M., & Espeland, W. N. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review*, *74*(1), 63–82. https://doi.org/10.1177/000312240907400104

Schiffer, Z. (2021, January 21). Instacart is firing every employee who voted to unionize. *The Verge*. https://www.theverge.com/2021/1/21/22242676/instacart-firing-every-union-employee-coronavirus-pandemic

Schneider, N. (2018). An Internet of ownership: Democratic design for the online economy. *The Sociological Review*, *66*(2), 320–340. https://doi.org/10.1177/0038026118758533

Scholz, T. (2017). *Uberworked and underpaid: How workers are disrupting the digital economy*. Polity.

Schomer, S. (2011, January 5). Zimride: Carpooling for college students. *Fast Company*. https://www.fastcompany.com/1702207/zimride-carpooling-college-students

Schor, J. (2020). *After the gig: How the sharing economy got hijacked and how to win it back*. University of California Press.

Schwandt, T. A., & Gates, E. F. (2018). Case study methodology. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (Fifth ed., pp. 600–630). Sage.

Schwartz, D. (2018). Embedded in the crowd: creative freelancers, crowdsourced work, and occupational community. *Work and Occupations*, *45*(3), 247–282. https://doi.org/10.1177/0730888418762263

Schwartz, J., Bohdal-Spiegelhoff, U., Gretczko, M., & Sloan, N. (2016). *The gig economy: Distraction or disruption?* https://www2.deloitte.com/global/en/insights/focus/human-capital-trends/2016/gig-economy-freelance-workforce.html

Schwellnus, C., Geva, A., Pak, M., & Veiel, R. (2019). *Gig economy platforms: Boon or Bane?* OECD Economics Department Working Papers. https://doi.org/10.1787/fdb0570b-en

Scott, J. (1990). *Domination and the arts of resistance: Hidden transcripts*. Yale University Press.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, *4*(2). https://doi.org/10.1177/2053951717738104

Sewell, G., & Barker, J. R. (2006). Coercion versus care: Using irony to make sense of organizational surveillance. *The Academy of Management Review*, *31*(4), 934–961. https://doi.org/10.2307/20159259

Shapiro, A. (2018). Between autonomy and control: Strategies of arbitrage in the "on-demand" economy. *New Media & Society*, *20*(8), 2954–2971. https://doi.org/10.1177/1461444817738236

Sharkey, A. J., & Bromley, P. (2015). Can ratings have indirect effects? Evidence from the organizational response to peers' environmental ratings. *American Sociological Review*, *80*(1), 63–91. https://doi.org/10.1177/0003122414559043

Sherman, R. (2011). Beyond interaction: Customer influence on housekeeping and room service work in hotels. *Work, Employment and Society*, *25*(1), 19–33. https://doi.org/10.1177/0950017010389240

Shestakofsky, B., & Kelkar, S. (2020). Making platforms work: Relationship labor and the management of publics. *Theory and Society*, *49*, 863–896. https://doi.org/10.1007/s11186-020-09407-z

Siddiqui, F. (2021, January 12). Gig workers and unions sue in California Supreme Court to have Prop 22 overturned. *Washington Post*. https://www.washingtonpost.com/technology/2021/01/12/prop22-uber-lawsuit/

Small, M. L. (2011). How to conduct a mixed methods study: Recent trends in a rapidly growing literature. *Annual Review of Sociology*, *37*, 57–86. https://doi.org/10.1146/annurev.soc.012809.102657

Smith, C. (2015). Continuity and change in labor process analysis forty years after *Labor and Monopoly Capital*. *Labor Studies Journal*, *40*(3), 222–242. https://doi.org/10.1177/0160449X15607154

Somerville, H. (2014, January 24). Startup Instacart hopes to dominate grocery delivery. *Mercury News*. https://www.mercurynews.com/2014/01/24/startup-instacart-hopes-to-dominate-grocery-delivery/

Soper, T. (2015, May 8). Lyft starting to show riders when they've been rated 5 stars by drivers. *GeekWire*. https://www.geekwire.com/2015/no-more-secrets-lyft-starting-to-show-riders-how-theyve-been-rated-by-drivers/

Srnicek, N. (2017). *Platform capitalism*. Polity.

Standing, G. (2011). *The precariat: The new dangerous class*. Bloomsbury.

Stanford, J. (2017). The resurgence of gig work: Historical and theoretical perspectives. *The Economic and Labour Relations Review*, *28*(3), 382–401. https://doi.org/10.1177/1035304617724303

Star, S. L., & Lampland, M. (2009). Reckoning with standards. In M. Lampland & S. L. Star (Eds.), *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life* (pp. 3–24). Cornell University Press.

Stark, D. (2009). *The sense of dissonance: Accounts of worth in economic life*. Princeton University Press.

Stark, D., & Pais, I. (2021). Algorithmic management in the platform economy. *Sociologica*, *14*(3), 47–72. https://doi.org/10.6092/issn.1971-8853/12221

Stark, L., & Levy, K. (2018). The surveillant consumer. *Media, Culture & Society*, *40*(8), 1202–1220. https://doi.org/10.1177/0163443718781985

Steinberg, R. J. (192). Gendered instructions: Cultural lag and gender bias in the Hay system of job evaluation. *Work and Occupations*, *19*(4), 387–423. https://doi.org/10.1177/0730888492019004004

Stewart, C. J., & Cash, W. B. (2003). *Interviewing: Principles and practices*. McGraw-Hill.

Stinson, L. (2016, November 15). Lyft ditches its iconic pink mustache to help you find your ride faster. *Wired*. https://www.wired.com/2016/11/lyft-ditches-iconic-pink-mustache-help-find-ride-faster/

Strathern, M. (2000). The tyranny of transparency. *British Educational Research Journal*, *26*(3), 309–321. https://doi.org/10.1080/713651562

Sullivan, C. (2009, July 29). Startup bets that social networking will spur carpool craze. *The New York Times*. https://archive.nytimes.com/www.nytimes.com/gwire/2009/07/29/29greenwire-startup-bets-that-social-networking-will-spur-36381.html

Sutherland, W., Jarrahi, MN. H., Dunn, M., & Nelson, S. B. (2020). Work precarity and gig literacies in online freelancing. *Work, Employment and Society*, *34*(3), 457–475. https://doi.org/10.1177/0950017019886511

Tadelis S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, *8*, 321–340. https://doi.org/10.1146/annurev-economics-080315-015325

TaskRabbit. (n.d.-a). *Register as a Tasker*. https://www.taskrabbit.com/become-a-tasker

TaskRabbit. (n.d.-b). *Welcome to the TaskRabbit Elite Facebook group*. https://taskrabbit1.typeform.com/to/zyh6kG?fbclid=IwAR3w12UNeofZw2P7yBvaa0Cb0DsRQwFEh6AUBO_lubr3kQDEk8AYqOEAcOQ
TaskRabbit. (n.d.-c). *How do I leave a review?* https://support.taskrabbit.com/hc/en-us/articles/213301766-How-Do-I-Leave-a-Review-
TaskRabbit. (2018). Fact Sheet. https://www.taskrabbit.com/press/Fact_Sheet.pdf
TaskRabbit. (2020). Fact sheet. https://www.taskrabbit.com/press/2020_T2_Fact_Sheet_US.pdf
Tassinari, A., & Maccarrone, V. (2020). Riders on the storm: Workplace solidarity among gig economy couriers in Italy and the UK. *Work, Employment, and Society, 34*(1), 35–54. https://doi.org/10.1177/0950017019862954
Taylor, F. W. (1911). *The principles of scientific management*. Harper & Brothers.
Taylor, P., & Bain, P. (2003). 'Subterranean worksick blues': Humour as subversion in two call centres. *Organization Studies, 24*(9), 1487–1509. https://doi.org/10.1177/0170840603249008
Tepper, F. (2015, August 17). Uber's new update gives food delivery as much as attention as transportation. *TechCrunch*. https://techcrunch.com/2015/08/17/ubers-new-update-gives-food-delivery-as-much-attention-as-transportation/
The Burrow. (2018, June 1). *Acceptance rate—What is it?* https://taskers.taskrabbit.com/2018/06/01/acceptance-rate-what-is-it/
The Fairwork Project. (2020). *The gig economy and Covid-19: Looking ahead*. https://fair.work/wp-content/uploads/sites/97/2020/09/COVID-19-Report-September-2020.pdf
The Fairwork Project. (2021). *Gig work principles.* https://fair.work/en/fw/principles/fairwork-principles-gig-work/
The Hutch by TaskRabbit. (2010, August 17). *Our runners are entrepreneurs*. https://blog.taskrabbit.com/our-runners-are-entrepreneurs
The Hutch by TaskRabbit. (2011, July 28). *Introducing the brand new TaskRabbit app for iPhone & iPod Touch!* https://blog.taskrabbit.com/introducing-the-brand-new-taskrabbit-app-for-iphone-ipod-touch
The Hutch by TaskRabbit. (2014a, March 10). *Introducing: The TaskRabbit Elite*. https://blog.taskrabbit.com/introducing-the-taskrabbit-elite
The Hutch by TaskRabbit. (2014b, June 17). *Unveiling the new TaskRabbit.* https://blog.taskrabbit.com/unveiling-the-new-taskrabbit
The Hutch by TaskRabbit. (2020, December 8). *A lookback at 2020: What we accomplished, together*. https://blog.taskrabbit.com/a-lookback-at-2020-what-we-accomplished-together
The University of Chicago Harris School of Public Policy., & the Associated Press-NORC Center for Public Affairs Research. (2020). *Pandemic disrupts demand for gig economy amidst safety concern*. https://harris.uchicago.edu/files/report_0.pdf
Thompson, A. I. (2015). Wrangling tips: Entrepreneurial manipulation in fast-food delivery. *Journal of Contemporary Ethnography, 44*(6), 737–765. https://doi.org/10.1177/0891241614550400

Ticona, J. (2015). Strategies of control: Workers' use of ICTs to shape knowledge and service work. *Information, Communication & Society*, *18*(5), 509–523. https://doi.org/10.1080/1369118X.2015.1012531

Ticona, J., & Mateescu, A. (2018). Trusted strangers: Carework platforms' cultural entrepreneurship in the on-demand economy. *New Media & Society*, *20*(11), 4384–4404. https://doi.org/10.1177/1461444818773727

Ticona, J., Mateescu, A., & Rosenblat, A. (2018). *Beyond disruption: How tech shapes labor across domestic work & ridehailing*. Data & Society. https://datasociety.net/library/beyond-disruption/

Tiku, N. (2015, April 28). Uber is bringing its food delivery service to New York City and Chicago. *The Verge*. https://www.theverge.com/2015/4/28/8505875/uber-ubereats-food-delivery-new-york-city-chicago

Toff, B., & Nielsen, R. K. (2018). "I just Google it": Folk theories of distributed discovery. *Journal of Communication*, *68*(3), 636–657. https://doi.org/10.1093/joc/jqy009

Truong, A. (2011, July 19). Don't want to run errands? Site lets you outsource them. *The Wall Street Journal*. https://www.wsj.com/articles/BL-DGB-22809

Tsotsis, A. (2012, July 2). Uber opens up platform to non-limo vehicles with "Uber X," service will be 35% less expensive. *TechCrunch*. https://techcrunch.com/2012/07/01/uber-opens-up-platform-to-non-limo-vehicles-with-uber-x-service-will-be-35-less-expensive/

Turow, J. (2017). *The aisles have eyes: How retailers track your shopping, strip your privacy, and define your power*. Yale University Press.

Uber Investor. (2020, November 5). *Uber announces results for third quarter 2020*. https://investor.uber.com/news-events/news/press-release-details/2020/Uber-Announces-Results-for-Third-Quarter-2020/

Uber Under the Hood. (2016, April 28). *Our approach to tipping*. Medium. https://medium.com/uber-under-the-hood/our-approach-to-tipping-aa0074c0fddc

Uber. (2014). *Your guide to the Uber rating system*. Retrieved from https://www.dropbox.com/s/ b0gnhj14o226y6l/The%20Rating%20System.pdf

Uber. (2015). *In response to the "sharing" economy: Issues facing platforms, participants, and regulators*. https://www.ftc.gov/system/files/documents/public_comments/2015/08/02049-96749.pdf

Uber. (2017). *You asked, and we're answering*. https://pages.et.uber.com/180-days/?_ga=2.45916900.1528381650.1614038996-99546731.1604443820

Uber Technologies, Inc. [Uber]. (2019a). Form S-1 registration statement. https://www.sec.gov/Archives/edgar/data/1543151/000119312519103850/d647752ds1.htm

Uber. (2019b, December 3). *To California drivers: Keeping you in the driver's seat*. Uber Blog. https://www.uber.com/blog/california/keeping-you-in-the-drivers-seat-1/

Uber. (2020, August 10). *A first step toward a new model for independent platform work*. Uber Newsroom. https://www.uber.com/en-ZA/newsroom/working-together-priorities/

UberCab. (2010, May 10). *UberCab full demo* [Video]. YouTube. https://www.youtube.com/watch?v=k3wk4kem54k&ab_channel=ubercab

US Government Accountability Office. (2015). *Contingent workforce: Size, characteristics, earnings, and benefits*. https://www.gao.gov/assets/gao-15-168r.pdf

Vallas, S., & Schor, J. B. (2020). What do platforms do? Understanding the gig economy. *Annual Review of Sociology*, *46*(1), 273–294. https://doi.org/10.1146/annurev-soc-121919-054857

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.

van Doorn, N. (2017). Platform labor: On the gendered and racialized exploitation of low-income service work in the 'on-demand' economy. *Information, Communication & Society*, *20*(6), 898–914. https://doi.org/10.1080/1369118X.2017.1294194

Veen, A., Barratt, T., & Goods, C. (2020). Platform-capital's 'App-etite' for control: A labour process analysis of food-delivery work in Australia. *Work, Employment and Society*, *34*(3), 388–406. https://doi.org/10.1177/0950017019836911

Volkin, M. (2020, May 27). *Why the gig economy will drive the future of employment*. *Forbes*. https://www.forbes.com/sites/forbescoachescouncil/2020/03/27/why-the-gig-economy-will-drive-the-future-of-employment/

Wachter, S., Mittelstadt, B. and Russell, C. (2018). Counterfactual explanations without opening the black box: automated decisions and the GDPR. *Harvard Journal of Law & Technology*, *31*(2), 841–887.

Wahl, C. (2020). *Organizational belonging in the new economy* [Doctoral dissertation, University of Pennsylvania]. ProQuest Dissertations Publishing. https://repository.upenn.edu/dissertations/AAI28087281/

Webb, A. (2016, December 9). *The "Uber for X" fad will pass because only Uber is Uber*. *Wired*. https://www.wired.com/2016/12/uber-x-fad-will-pass-uber-uber/

Wedlin, L. (2007). The role of rankings in codifying a business school template: Classifications, diffusion and mediated isomorphism in organizational fields. *European Management Review*, *4*(1), 24–39. https://doi.org/10.1057/palgrave.emr.1500073

Wharton, A. S. (2015). Interactive service work. In S. Edgell, H. Gottfried, & E. Granter (Eds.), *The Sage handbook of the sociology of work and employment* (pp. 329–345). Sage.

Wired. (2018, October 12). WIRED25: The future of work with Stacy Brown-Philpot, CEO of TaskRabbit [Video]. *Wired*. https://www.wired.com/video/watch/wired25-the-future-of-work-with-stacy-brown-philpot-ceo-of-taskrabbit

Wolfe, A. (2017, August 4). TaskRabbit CEO Stacy Brown-Philpot's task: Expansion. *The Wall Street Journal*. https://www.wsj.com/articles/taskrabbit-ceo-stacy-brown-philpots-task-expansion-1501875939

Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. *Work, Employment and Society, 33*(1), 56–75. https://doi.org/10.1177/0950017018785616

Wood, A. J., Lehdonvirta, V., & Graham, M. (2018). Workers of the Internet unite? Online freelancer organisation among remote gig economy workers in six Asian and African countries. *New Technology, Work and Employment, 33*(2), 95–112. https://doi.org/10.1111/ntwe.12112

Woodcock, J. (2017). *Working the phones: Control and resistance in call centres*. Pluto Press.

Woodcock, J., & Graham, M. (2020). *The gig economy: A critical introduction*. Polity.

Working Washington. (2019, January 17). *Instacart: Here's our 22 cents — no more tip theft, low pay, and black-box pay algorithms*. Medium. https://medium.com/@workingwa/instacart-heres-our-22-cents-no-more-tip-theft-low-pay-and-black-box-pay-algorithms-8ff1d7c6b66

Y Combinator. (2014, June 20). *Apoorva Mehta at startup school NY 2014* [Video]. YouTube. https://www.youtube.com/watch?v=wkmR7TYUt_c&ab_channel=YCombinator

Y Combinator. (2018, October 17). *Q&A with Apoorva Mehta, Founder & CEO, Instacart*. https://blog.ycombinator.com/qa-with-apoorva-mehta-founder-ceo-instacart/

Yeo, L. (2021, February 18). Uber vs. Lyft: Who's tops in the battle of U.S. rideshare companies. *Second Measure*. https://secondmeasure.com/datapoints/rideshare-industry-overview/

Yin, R. K. (2012). *Applications of case study research*. Sage.

Ytre-Arne, B., & Moe, H. (2020). Folk theories of algorithms: Understanding digital irritation. *Media, Culture & Society*. Advance online publication. https://doi.org/10.1177/0163443720972314

Ziewitz, M. (2016). Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values*, *41*(1), 3–16. https://doi.org/10.1177/0162243915608948

Ziewitz, M. (2019). Rethinking gaming: The ethical work of optimization in web search engines. *Social Studies of Science, 49*(5), 707–731. https://doi.org/10.1177/0306312719865607

APPENDICES

Appendix 1-1. The Interview Guide I Used for the Project about Uber Drivers' Performance of Emotional Labor

1. How long have you been working as an Uber driver?
2. How did you become an Uber driver?
3. How would you describe a day at your job?
4. Do you use any mobile apps or technologies during your work? If so, how do you use these technologies?
5. How do you use the Uber app?
6. What do you think of the Uber apps (May probe: what about the surge pricing, the rating system, and other features?)
7. How do you use the Uber rating system? Do you have any strategies to get a good rating?
8. How do you use the rating system to evaluate the riders?
9. Tell me about problems you have encountered at work. What is the hardest thing about your job? What did you do?
10. What is the most rewarding thing about your job?
11. Describe how you usually interact with the riders. Could you also describe your relationship with the rider?
12. How do you decide whether (or how much) you talk with the rider during the ride?
13. Where do you get advice on how to interact with the riders?
14. Some people have told me that they are concerned about the Uber ratings. Is that something you think about?
15. What are the required skills for an Uber driver?
16. What would you change about your job if you could do it?
17. Is there anything else you would like to add?

Appendix 1-2. In-Person Recruitment Message During a Ride

Hi! I'm Oliver, a PhD student in Communication at Cornell. I'm studying how Uber drivers use mobile technologies and interact with riders. I was wondering if I could interview you over Skype or phone when you are free. It should take about 15-20 minutes. It is entirely voluntary and nothing will be attributed to you. And you can skip any questions you want. You should know that your participation will also not affect the compensation for your ride service. Does that sound ok? Shall we set up a time to do the interview? If you are unable or unwilling to participate in this study, do you know someone who might be willing to talk to me about their Uber driving experience? Thank you very much!

Appendix 1-3. Sample Recruitment Email for Uber Drivers who Shared Content on YouTube or Their Blogs

Dear *Name of the blogger/ owners of online forum or social media group*,

My name is Oliver Ngai Keung Chan, a PhD student in the Department of Communication at Cornell University. I am conducting a study on how Uber drivers use mobile technologies and interact with riders. Because *your blog/ online forum/ social media group (either one of these terms, depending on the platform)* is about Uber drivers, I was wondering if I could interview you over Skype or phone to know more about Uber drivers' experience. The interview should take about 30 to 45 minutes. If you are unable or unwilling to participate in this study, do you know someone who might be willing to talk to me about their Uber driving experience?

Your participation is entirely voluntary. If you are interested in participating in this study or have any questions about the study, please feel me to contact me by phone (phone number removed) or email (nc478@cornell.edu). I look forward to hearing from you. Thank you so much!

Sincerely,
Oliver Ngai Keung Chan

Appendix 1-4. Screenshot of the Recruitment Material I used on Twitter

Interview Participants Needed

Hello Everyone! I'm Oliver Ngai Keung Chan, a Ph.D. candidate in the Department of Communication at Cornell University. I'm currently working on my dissertation research about gig workers' experiences on labor platforms. I'm looking to chat with **Uber/Lyft drivers, TaskRabbit Taskers,** or **people who work for platform co-operatives** for my research project.

I'm particularly interested in learning more about how workers use technologies and metrics (e.g., online ratings) as well as interact with their customers. The interview should take about **45 minutes to 1 hour** over **Skype or phone** at your convenience. After the interview, I will provide you a $20 gift card to compensate for your time and insights.

If you are interested in participating in the study, or would like to know more about it, please feel free to DM me (@ngaikeungchan) on Twitter or email me (nc478@cornell.edu) for more information. I would love to hear from you!

Appendix 1-5. The Interview Guide for Uber Driver/Bloggers

Note: The ordering of the interview questions was slightly changed depending on what videos or content the interviewees had already posted online.

A. Questions about Drivers' Uses of Mobile Technologies and Emotional Labor

1. How long have you been working as an Uber driver?
2. How did you become an Uber driver?
3. How would you describe a day at your job?
4. Do you use any mobile apps or technologies during your work? If so, how do you use these technologies?
5. How do you use the Uber app?
6. What do you think of the Uber apps (May probe: what about the surge pricing, the rating system, and other features?)
7. How do you use the Uber rating system? Do you have any strategies to get a good rating?
8. How do you use the rating system to evaluate the riders?
9. Describe how you usually interact with the riders.
10. How do you decide whether (or how much) you talk with the rider during the ride?
11. What is the most rewarding thing about your job?
12. Tell me about problems you have encountered at work. What is the hardest thing about your job? What did you do?
13. Where do you get advice on how to interact with the riders?
14. What would you change about your job if you could do it?

B. Questions about Performance of Expertise and Self-Branding Activities

15. How did you become an Uber blogger/ YouTuber?
16. Could you describe the usual process of recording and posting a video on your YouTube channel (or writing, editing, and posting an article online)?
17. How often do you record and post a video on your YouTube channel (or an article online)?
18. Do you pay attention to number of views, likes, and comments on your videos/ articles?
19. What are the kinds of videos/ articles that attract a lot of views, in your opinion?
20. What are the strategies to promote your YouTube channel/ website, in your opinion?
21. Who is your audience? How do you interact with your audience?
22. Do you interact with other Uber bloggers online?
23. What is the best part of sharing videos and posts online?
24. What is the worst/ least favorite part of sharing videos and posts online?
25. How do you balance your work as an Uber driver and your blog/ YouTube channel?
26. Some people have told me about the idea of creating a personal brand. What does personal branding mean to you?

C. Closing Question

27. Is there anything else about your work that I haven't asked that you think would be useful to share?

Appendix 1-6. The Interview Guide for the Dissertation Project

Note: The interview questions were slightly modified for workers who worked on different platforms, depending on platform interfaces and uses of the metrics.

1. How long have you been working as a Lyft driver/ Uber driver/ TaskRabbit tasker/ Instacart shopper/ DoorDash Dasher/ Uber Eats driver?
2. How did you become a Lyft driver/ Uber driver/ TaskRabbit tasker/ Instacart shopper/ DoorDash Dasher/ Uber Eats driver?
3. How would you describe a day at your job?
4. How many hours do you work for Lyft/ Uber/ TaskRabbit/ Instacart/ DoorDash/ Uber Eats in a week?
5. How do you use the Lyft/ Uber/ TaskRabbit/ Instacart/ DoorDash/ Uber Eats app? Could you please walk me through the app? (Prompt: Ask participants to go over each step, from taking a request to view their metrics).
6. [Only for workers on Uber/ Lyft/ TaskRabbit] How do you manage your online profile?
7. Could you please pull up your metrics and describe to me what you see?
8. How do you use these metrics in your daily work?
9. Have you ever received a negative review or rating? How does it feel?
10. Some people have told me that they are concerned about their metrics. Is that something you think about?
11. Do you participate in Uber Pro/ Lyft rewards program/ the TaskRabbit Elite/ Top Dasher program? How do you feel about this program?
12. Describe how you usually interact with the riders/ clients/ customers. Could you also describe your relationship with them?
13. Where do you get advice on how to interact with the riders/ clients/ customers?
14. How do you usually communicate with the company you work for?
15. In your opinion, what is good service?
16. In your opinion, what are the required skills for a Lyft driver/ Uber driver/ TaskRabbit tasker/ Instacart shopper/ DoorDash Dasher/ Uber Eats driver?
17. What is the most rewarding thing about your job?
18. Tell me about problems you have encountered at work. What is the hardest thing about your job? What did you do?
19. What would you change about your job if you could do it?
20. What are your future career goals?
21. Is there anything else you would like to add?

Appendix 1-7. Sample Message of the Reddit Advertisements for Participant Recruitment

PROMOTED · Posted by u/Oliver_ChanNK 2 months ago

0

Looking to interview Taskers for a paid research project (Remote interview)

Hi r/TaskRabbit,

My name is Oliver Chan, a Ph.D. candidate in the Department of Communication at Cornell University. I am currently working on my dissertation research about gig workers' experiences on labor platforms. I'm looking to interview TaskRabbit Taskers for my research project. I'm particularly interested in learning more about how works use technologies and metrics (e.g., ratings) as well as interact with their customers. The interview is completely confidential will take about 45 minutes virtually or via phone at your convenience. After the interview, I will provide you with a $20 gift card to compensate for your time and insights.

If you are interested in participating in the study, or would like to know more about it, please feel free to DM me or email me (nc478@cornell.edu). Thank you so much in advance!

1 Share ••• 50% Upvoted

Comment as Oliver_ChanNK

Appendix 1-8. Sample Setting of the Reddit Advertisements for Participant Recruitment

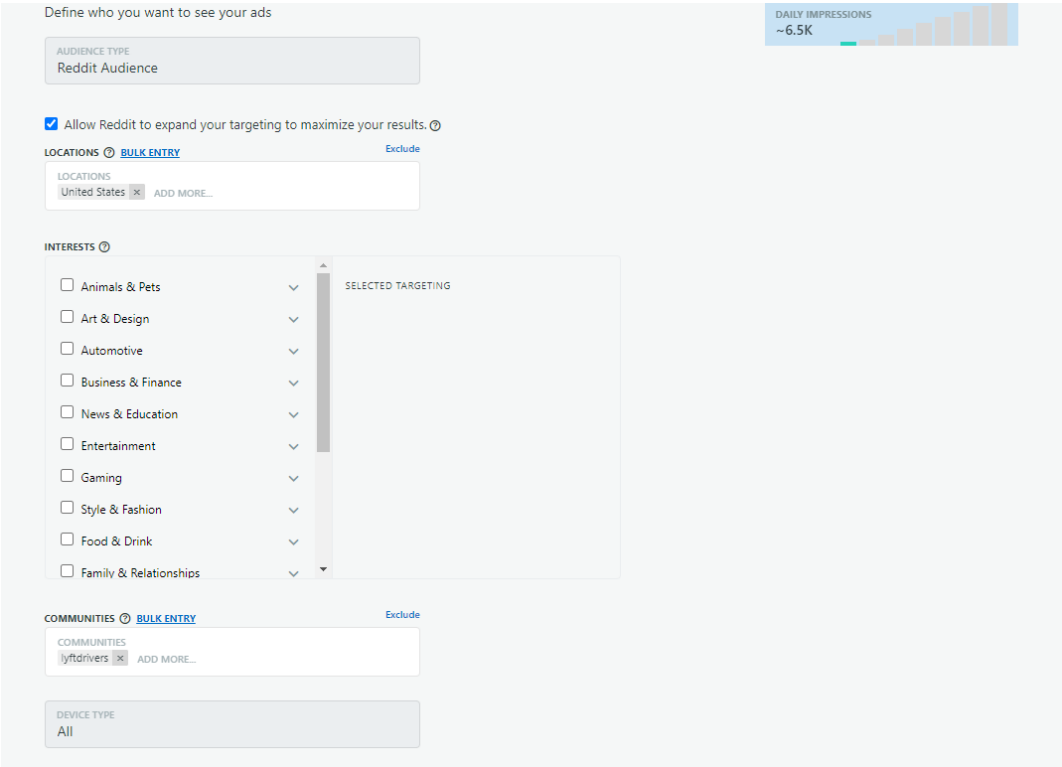


Table 3-1. List of Interviewees who Worked for Ride-hailing Platforms

Pseudonym	Platforms	Sex	Working location	Length of working on the platform	Average working hours per week
Lincoln	Uber and Lyft	Male	Syracuse	6 months	Part-time (usually worked on Saturday nights)
Timothy	Uber and Lyft	Male	Syracuse	6 months	Part-time (varied week by week)
Daniel	Uber	Male	Ithaca	3 months	Usually worked 5 to 7 days (from 9 pm to 6 am)
Noah	Uber	Male	Los Angeles	1 year and 10 months	Full-time (12-hour shift on a typical day)
Clayton	Uber and Lyft	Male	Los Angeles	Almost 3 years	Unstable (4-hour shift on a typical day)
Austin	Uber and Lyft	Male	Chicago	2 years and 2 months	26 to 32 hours
Cohen	Uber and Lyft	Male	Los Angeles	4 years	NA
Noel	Uber	Male	Louisville	2 years	Over 40 hours
Howell	Uber and Lyft	Male	San Mateo	About 6 years	10 to 20 hours
Rodrigo	Uber and Lyft	Male	Charlotte	1 year and 10 months	40 hours
Charles*	Uber, Lyft, and Uber Eats	Male	Miami	2 years (Uber and Lyft) and 1 year (Uber Eats)	NA
Lyndsey	Uber and Lyft	Female	Baltimore	1 year and 9 months	10 hours
Alexis	Uber and Lyft	Female	Chicago	1 year	Depending on Uber's promotions
Sadie	Uber and Lyft	Female	Monterey	1 year	NA
Jeremy	Uber	Male	San Diego	4 months	15 to 20 hours

Dominic	Uber and Lyft	Male	Georgia	5 years	25 to 30 hours
Brandon	Uber and Lyft	Male	Chicago	5 years	30 hours (Uber) and 2 hours (Lyft)
Eddy	Uber and Lyft	Male	Washington, DC	3 months in 2014 (Uber) and 11 months between 2017 and 2018 (Lyft)	30 to 35 hours
Adrian	Uber and Lyft	Male	Los Angeles	3.5 years	72 hours
Kimberly	Uber	Female	Los Angeles	1 year	15 hours
Osmar	Lyft	Male	Chicago	2 years	40 hours

Note. At the time when I interviewed Charles, my project primarily focused on Uber drivers' work experiences. "NA" means that I forgot to ask about average working hours during the interviews.

Table 3-2. List of Interviewees who worked for TaskRabbit

Pseudonym	Sex	Working location	Length of working on the platform	Average working hours per week
Kuzma	Male	Cincinnati	1 year and 8 months	45 hours
Kooper	Male	Atlanta	9 years	20 to 25 hours (sometimes up to 40 hours)
Zoey	Male	Los Angeles	7 years	Unstable (depending on her schedule and what tasks she was hired)
Jeffery	Male	East Coast	3 years	44 hours
Easton	Male	Dallas	2 years	6 hours
Renee	Female	Denver	2 years	20 to 30 hours
Maverick	Male	Dallas–Fort Worth Area	6 months	15 hours
Scarlett	Female	Los Angeles	4 years	10 hours
Neyland	Male	Denver	1 year	10 hours
Audrey	Female	Portland	1 year	20 hours
Murdock*	Male	San Diego	6 months (TaskRabbit) and 1 year and 8 months (Uber Eats)	30-40 hours (TaskRabbit) & Uber Eats (16 hours)

Note. Murdock worked for TaskRabbit and Uber Eats, but he primarily used TaskRabbit.

Table 3-3. List of Interviewees who worked for Delivery Platforms

Pseudonym	Platforms	Sex	Working location	Length of working on the platform	Average working hours per week
Collin	Uber Eats	Male	Massachusetts	1 year	15 to 20 hours
Joseph	Instacart (full-service shopper)	Male	Chicago	4 months	10 hours
Adele	Instacart (in-store shopper)	Female	Los Angeles	1 year and 2 months	8 hours (during the pandemic) & 25 hours (before the pandemic)
Derek	Instacart (full-service shopper)	Male	Maryland	7 months	60 to 70 hours
Connor	Uber Eats and DoorDash	Male	Massachusetts	2.5 years	40 hours
Clara	DoorDash	Female	Massachusetts	2.5 years	6 hours
Carter	DoorDash	Male	Massachusetts	2.5 years	40 hours
Melvin	DoorDash	Male	Pennsylvania	3 months	4 to 6 hours
Sean	DoorDash	Male	New York	1.5 years	20 to 25 hours
Nelson	DoorDash	Male	Central Indiana	6 months	5 to 10 hours
Shawn	DoorDash	Male	Olive Branch, Mississippi	3 months	10 hours
Jackson	Uber Eats	Male	Atlanta	3.5 years	25 to 30 hours (sometimes up to 40 hours)
Roger	Uber Eats and DoorDash	Male	Arlington	1 year	20 hours
Sofia	DoorDash	Female	Portland	2 to 3 months	15 minutes to 3 hours
Dylan	Uber Eats	Male	Washington, DC	6 months	12 hours
Lorenzo	Uber Eats and DoorDash	Male	Southwest Virginia	2 years	35 to 45 hours
Harvey	Uber Eats and DoorDash	Male	Columbus	4 months	12 hours

Ronnie*	DoorDash and Lyft	Male	Las Cruces, New Mexico	11 months (DoorDash) and Lyft (4 months)	20 to 25 hours (DoorDash) & 30 to 35 hours (Lyft)

Note. Ronne usually turned on both DoorDash and Lyft apps to work at the same time. He worked on DoorDash for a longer time period than Lyft, but he just started spending more time on Lyft at the time I interviewed him.

Table 4. Overview of Platforms' Worker-related Metrics

	Types of metrics	Methods of calculating customer ratings	Dispute Process
Uber	5-star customer ratings, acceptance rates, cancellation rates, and Uber Pro points (for unlocking tiers in the rewards program)	Average star ratings based on the last 500 rated rides	No individual ratings can be removed upon request
Lyft	5-star customer ratings, acceptance rates, cancellation rates, Qualifying points, and Lyft Driving Points (the latter two are for unlocking tiers in the rewards program)	Average star ratings based on the last 100 rated rides. If a passenger does not leave a rating, drivers will automatically get a 5-star rating. The single lowest rating will be removed for every 100 rides.	Drivers can contact Lyft, if they believe that a rating is unfair
TaskRabbit	5-star customer ratings, response rates, acceptance rates, reliability rates	Percentage of positive reviews (3-5 stars) based on all rated Tasks, but individual star ratings remain visible to prospective clients	Taskers can contact TaskRabbit to dispute an unfair rating, but TaskRabbit discourages Taskers to contact their clients about a negative review or rating
DoorDash	5-star customer ratings, restaurant ratings (thumbs up or thumbs down), acceptance rates, completion rates, and on-time/early rates	Average star ratings based on the last 100 rated delivery orders	DoorDash will automatically remove unfair ratings in certain situations (e.g., long wait times at restaurants, extreme weather conditions, and system-wide outages, etc.)
Uber Eats	Customer and restaurant ratings	Percentage average of thumbs up	No individual ratings can be removed upon

	(thumbs up or thumbs down), acceptance rates, cancellation rates, and Uber Pro points	ratings based on the last 100 rated delivery orders	request
Instacart	5-star customer ratings, customer feedback (communication, replacing items, quality of items, and finding items; thumbs up or thumbs down), cancellation rates, and speed metrics	Average star ratings based on the last 100 rated delivery orders. For every 100 orders, the lowest rating is automatically removed.	Shoppers can contact Instacart if they believe that a rating is unfair