Platform Design and Operations in the Last-Mile of the Supply Chain

by

Lina Wang

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved June 2021 by the
Graduate Supervisory Committee:

Elliot Rabinovich, Chair
Timothy Richards
Scott Webster
Harish Guda

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

Platform business models have become pervasive in many aspects of the economy, particularly in the areas experiencing rapid growth such as retailing (e.g., Amazon and eBay) and last-mile transportation (e.g., Instacart and Amazon Flex). The popularity of platform business models is, in part, due to the asset-light prospect which allows businesses to maintain flexibility while scaling up their operations. Yet, this ease of growth may not necessarily be conducive to viable outcomes. Because scalability in a platform depends on the intermediary's role it plays in facilitating matching between users on each side of the platform, the efficiency of matching could be eroded as growth increases search frictions and matching costs. This phenomenon is demonstrated in recent studies on platform growth (e.g. Fradkin, 2017; Lian and Van Ryzin, 2021; Li and Netessine, 2020).

To sustain scalability during growth, platforms must rely on effective platform design to mitigate challenges arising in facilitating efficient matching. Market design differs in its focus between retail and last-mile transportation platforms. In retail platforms, platform design's emphasis is on helping consumers navigate through a variety of product offerings to match their needs while connecting vendors to a large consumer base (Dinerstein *et al.*, 2018; Bimpikis *et al.*, 2020). Because these platforms exist to manage two-sided demand, scalability depends on the realization of indirect network economies where benefits for users to participate on the platforms are commensurate with the size of users on the other side (Parker and Van Alstyne, 2005; Armstrong, 2006; Rysman, 2009). Thus, platform design plays a critical role in the realization of indirect network economies on retail platforms.

Last-mile transportation platforms manage independent drivers on one side and retailers on the other, both parties holding flexibility in switching between platforms. High demand for independent drivers along with their flexibility in work participation

induces platforms to use subsidies to incentivize retention. This leads to short-term improvements in retention at the expense of significant increases in platforms' compensation costs. Acute challenges to driver retention call for effective compensation strategies to better coordinate labor participation from these drivers (Nikzad, 2017; Liu *et al.*, 2019; Guda and Subramanian, 2019). In addition to driver turnover, retailers' withdrawal can undermine the operating efficiency of last-mile transportation platforms (Borsenberger *et al.*, 2018). This dissertation studies platforms' scalability and operational challenges faced by platforms in the growth.

DEDICATION

For my family

ACKNOWLEDGMENTS

Undertaking this Ph.D. has been a truly life-changing experience for me. I am so grateful for the support and guidance that I received from an extraordinary group of people who are patient, generous with their time as well as their advice.

First and most important, I would like to thank my advisor, Dr. Elliot Rabinovich. I am so grateful for his guidance and commitment from the first day of my doctorate studies. He invested countless hours in patiently coaching me on how to be a scholar and guiding me through the journey from idea generating to paper crafting to journal review processes. I could not have remotely imagined that I could be the researcher I am today without the care and effort he consistently exerted to nurture my growth in the past five years.

I also really appreciate all of the guidance from the rest of my committee members. Thank you to Dr. Timothy Richards, who introduced me to a series of rigorous empirical methodologies valued by economists. Thank you in particular for taking your time to patiently answer any questions I encountered in research, big and small. In addition, thank you to Dr. Scott Webster for being a part of my dissertation committee. Dr. Webster showed me how to connect empirical work with analytical frameworks and provided invaluable perspectives on the dissertation. Last but not least, I would like to thank Dr. Harish Guda, who always made himself available to me and gave great advice not only on my dissertation but also on my academic career. Thank you for your inspiration on this intellectual journey.

I must acknowledge the Department of Supply Chain Management for funding and supporting my studies. My studies would not have been possible without the immense patience and help of all the faculty who have furthered my education both in and outside the classroom. Thank you to Dr. Choi, Dr. Kull, and Dr. Yin for teaching the doctoral seminars that provided me with the foundational field knowledge that I was able to build on. I am also grateful to Dr. Carter, Dr. Li, Dr. Polyviou, Dr. Rogers, and Dr. Wiedmer who took the time to review my summer papers and helped me make invaluable improvements. I would also like to thank Eddie Davila for showing me how to teach effectively. I appreciate Dr. G, the chair of our department, for all of his encouragement. I thank Christa Thompson, Lydia Chang, and Angela Lentino for their excellent administrative assistance. Finally, I was so lucky to have had excellent company working with Ph.D. students who acted as intellectual, philosophical, and moral guides helping me navigate through the ebb and flow of my studies.

I must also acknowledge TForce Logistics for supporting this dissertation. A big thank you to Guy Farthing who provided me with access to data and took the time to provide valuable managerial insights on operations of last-mile delivery.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table | Page

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

SCALABILITY IN PLATFORMS FOR LOCAL GROCERIES: AN EXAMINATION OF INDIRECT NETWORK ECONOMIES

Abstract

Despite a significant rise in consumer interest in local foods, supply constraints limit access to these products in many markets. Online platforms for local foods may help solve these constraints. However, to our knowledge, there is no empirical research on the economic viability of these platforms. We study this problem by analyzing a two-sided platform subject to indirect network effects. If present, these effects will create a virtuous cycle where consumers' demand for products sold through the platform rises in the number of vendors and suppliers' demand for product distribution through the platform increases in consumer demand. In the case of our study's platform, analyses reveal the existence of indirect network effects, as consumers prefer a variety of local vendors and vendors derive greater surplus from greater consumer demand. Therefore, platforms like the one we analyze may serve as viable alternatives for the commercialization of local foods, and could provide greater access to these products. Importantly, however, our analysis also reveals the existence of not only non-linearities in the strength of indirect network effects, but also non-monotonic effects. Non-monotonicity derives from consumers' attraction to the platform marginally decreasing in the number of local vendors and from the existence of marginally increasing costs as more of these vendors join in. As a result, indirect network economies are subject to a cap imposed by the number of vendors participating in these platforms. Through

counterfactual simulations, we evaluate the magnitude of this constraint and offer recommendations on how to minimize its impact.

1 Introduction

Increasing demand for locally-produced foods continues to place new strains on food supply chains traditionally designed to widely distribute mass-produced national brands (Voight, 2013; Hesterman and Horan, 2017). Local foods appeal to consumers due to their perceived quality and freshness (Thilmany *et al.*, 2008; Bond *et al.*, 2009) and a sense that they are easier to trace to their source, are more environmentally sustainable, and contribute to the local economy in a positive way (Carpio and Isengildina-Massa, 2009; Toler *et al.*, 2009). Nevertheless, despite this rise in demand, supply constraints limit access to local foods in many markets (Tippins *et al.*, 2002). Historically, local foods were first available only through direct-to-consumer channels (i.e., specialty stores, farmers' markets, restaurants, and roadside stands), but direct channels are not available in all markets (Hardesty, 2008). Once consumers' demand for local foods became apparent, traditional grocery chains, such as Walmart (Bloom and Hinrichs, 2017), scrambled to meet this demand by stocking more local products. However, these chains typically stock only a limited product selection, tend to negotiate prices and inventory-management terms that are not economically viable for smaller, local producers, and require these suppliers to navigate multiple layers of intermediaries and distribution centers to supply different stores (Swanson, 2013; Gunders and Bloom, 2017; Tongarlak *et al.*, 2017; Zuurbier, 1999; Mena *et al.*, 2011). Consequently, many local-vendor relationships soon fail; witness the mass de-listing of local food products by Tesco in 2015 (Leyland, 2015; Willis *et al.*, 2016). In this paper, we consider an alternative, Internet-based platform model that bypasses traditional retailers, and leverages the same fundamental economic forces that drive similar models such as those involving mobile apps, gaming, music, and video.

This model differs from others also based on the Internet and used by more established online retailers, such as Fresh Direct. Although these retailers are also able to source fresh foods directly from small specialty vendors using disintermediated supply chains (Cattani *et al.*, 2007; Richards *et al.*, 2017; Stewart *et al.*, 2018), they have yet to be able to solve the geographic disconnect that exists between vendors of local products, and the critical mass of the consuming market, much of which lives in major urban centers, such as New York City (Low and Vogel, 2011). Therefore, although online grocery retailing models can be effective in disintermediating food supply chains, they do not offer complete solutions to consumers' lack of direct access to local food vendors located in the same areas where consumers live.

Emerging, second-generation models based on Internet platforms may provide a solution to this challenge (Horst *et al.*, 2011; Matson *et al.*, 2013; Richards *et al.*, 2017). In these models, the platform typically exists solely to connect producers of local foods to consumers in a geographically proximate area. Serving as an aggregator for local foods, the platform provides a means by which producers can distribute their goods in their own local markets and consumers in those markets can obtain easy access to these products through direct delivery. Yet, despite a growing interest in such Internet-enabled food hubs, no research has examined their economic viability. Although there are a number of platforms that are attempting to move fresh food directly from local producers to consumers, e.g., Farmhouse (farmhousedelivery.com) and Farm Runners (Farmrunners.com), these firms have yet to generate the kind of growth in their local supply base that would suggest they will be a real alternative to traditional grocery channels for local foods. [1] In this paper, we provide an empirical analysis of transaction data from one Internet-based local food platform in order to test whether or not the underlying conditions for long-term scalability exist. In doing so, we not only provide new insights as to whether this type of second-generation model is economically viable, but also examine whether it can grow to include a substantial number of local vendors in the market. That is, we establish whether the viability of growth in this type of platform is constrained in the number of vendors participating in the platform.

Our study focuses on a platform that uses a model commonly observed in the industry. In this model, participating consumers first order products from an array of local vendors and pay prices set by the platform. Upon receiving the orders, the platform then uses its own vehicles and employees to pick up the items from the vendors and deliver them to consumers. Therefore, the primary benefit of a platform such as this is that it reduces transaction costs between consumers and vendors, and provides these parties low-cost access to local grocery inventories and distribution services (i.e., product pickup from the vendors and delivery to end consumers), respectively. That is, the economic viability of the platform depends, in part, on reducing frictions (i.e., transaction costs and search costs) involved in matching consumer demand for items in local vendors' inventories with local vendors' demand for the distribution of these inventories to consumers. Perhaps more important, however, are the more subtle *indirect network effects.*

The platform in our study is considered to be a "pure platform" in the sense that it exists to manage two-sided demand: Consumers demand a variety of products from vendors, and vendors demand distribution to the largest possible number of consumers.

[1] Although online platforms in these markets have been operational for 7 years on average, most of them (70%) include fewer than 40 local vendors in their supply base (USDA Local Food Directories, 2020).

Therefore, viability depends on the realization of indirect network economies between consumers on one side of the platform and vendors on the other (Rochet and Tirole, 2006; Rysman, 2009). Because the platform is two-sided, the strength of any indirect network effects will rise in (1) the economic gains available for vendors from using the platform to distribute their inventory to satisfy greater consumer demand, and (2) the utility consumers receive from buying from a larger number of vendors and, hence, a greater variety of local items to choose from. Indirect network effects create a virtuous cycle in which demand from consumers supports a larger number of vendors on the platform, and more vendors will attract greater consumer demand (Tucker and Zhang, 2010). Stronger indirect network effects, therefore, generate more surplus for consumers participating on the platform and for vendors that choose to distribute their products through the platform.

By studying this phenomenon, we contribute to the emerging operations management (OM) literature on platforms (see Benjaafar and Hu, 2020 and Chen *et al.*, 2020 for a review) and to research examining the role of platforms in the management of supply chains (e.g., Parker and Anderson, 2002). This is because the platform we analyze must compete with grocery stores, specialty stores, and farmers' markets for supply and demand within the same geographical market. Therefore, our evaluation of indirect network economies in this setting is based on the surplus available to the platform in equilibrium. Among recent studies that estimate indirect network effects that are most similar to ours (Chu and Manchanda, 2016; Li and Netessine, 2020; Zhou *et al.*, 2020), none allow for the observation that indirect network effects arise as a result of an equilibrium in the interaction between buyers and sellers on the platform.

This distinction is important. For example, while Chu and Manchanda (2016) explain both buyer and seller behavior independently on a marketplace in terms of utility-based models, the marketplace does not choose prices and variety based on its profit-maximization behavior conditional on consumer demand. Our approach, on the other hand, ensures that the indirect network effects we estimate are fully consistent with each set of agents optimizing their respective objective functions based on their expectations of the other side's decisions. Essentially, the model by Chu and Manchanda (2016) is a reduced-form explanation of indirect network effects, while ours accounts explicitly for the endogeneity of each side.

At the core of our study is an empirical model that is similar in spirit to others used in analyses of indirect network effects in marketing and economics by Nair *et al.* (2004), Kaiser and Wright (2006), and Richards and Hamilton (2018), but also extends these analyses in two ways. First, our evaluation of the strength of the indirect network effects depends on the competitive conduct of the platform in equilibrium, since monopolistic conduct could effectively neutralize indirect network effects by internalizing them (Weyl, 2010). In our model, the platform's pricing decisions are conditioned on the equilibrium between consumer demand for local foods and supplier demand for distribution, and on the equilibrium responses by the platform's competitors. As the platform's conduct deviates from perfect competition, then its pricing decisions will increasingly absorb the indirect network effects as profit to the platform owner (Weyl, 2010). The results from our model show that the platform's conduct does deviate from perfect competition, absorbing some of the available indirect network effects.

Second, our study considers non-linearities and non-monotonicity in the extent to which consumer demand rises in the number of vendors on the platform and vice versa. While prior work on indirect network effects has documented how two-sided demand can generate surplus for different platforms (e.g., Rochet and Tirole, 2003; Rysman, 2004; Parker and Van Alstyne, 2005), it has assumed that indirect network effects are either linear or nonlinear but monotonic. We, on the other hand, consider a more realistic case in which a platform's indirect network effects may not be inherently unlimited. That is, we analyze whether indirect network effects continue unabated as the platform's scale increases by considering nonlinear and non-monotonic network effects. This is another contribution our paper makes to the literature. Although prior studies contribute theoretical models that account for these conditions (Halaburda *et al.*, 2017), we are not aware of research that has modeled them empirically.

Our analysis of indirect network economies is structural in the sense that we evaluate the available platform surplus for participating vendors as a nonlinear function of consumer demand and the attraction of consumers to the platform as a nonlinear function of the number of vendors. The use of this structural approach represents another contribution of our study to the body of research on platforms in the OM literature, which typically uses empirical methods that rely on quasi-randomization and exogenous sources of variation in the data to characterize nonlinear indirect network economies (e.g., Li and Netessine, 2020). Our structural approach offers the opportunity to understand *why* nonlinear indirect network economies occur in the first place.

We find that there are substantial indirect network effects in an Internet-based local food platform such as the one in our study. That is, we find that consumer demand increases in the number of vendors participating in the platform, and vice versa. Therefore, a platform like the one we analyze may constitute a viable alternative for the commercialization of local grocery food to consumers and could offer a solution to consumers' limited access to local food vendors through supply chains involving grocery retailers, specialty stores, and farmers' markets.

However, we also find that the indirect network effects in the platform are nonlinear in the number of participating vendors and that these effects are non-monotonic. Thus, the network effects we document do not increase without bound and there is an economic limit to the platform's scale. First, we observe that having too many vendors can erode the utility consumers derive from the platform. As a result, indirect network economies at the platform will be subject to negative *cross-side externalities* when the number of participating vendors grows too large. Second, we find that indirect network effects are subject to negative *same-side externalities* on the supply side because increases in the number of participating vendors can yield marginally increasing costs in fulfilling consumer demand at the platform. [2]

Alleviating these negative externalities will allow the platform to give more local vendors an opportunity to market their inventories outside traditional local channels and grocery chains. Therefore, we conclude our paper by carrying out two different counterfactual simulations directed at achieving this goal. The first simulation addresses the negative *cross-side externalities* by increasing consumers' preferences for local food vendors through improvements in the platform's search and recommendations capabilities. The second simulation addresses the negative *same-side externalities* by offsetting the marginal costs at the platform through the use of government subsidies promoting the provision of local grocery foods. These counterfactual simulations allow us to identify the number of suppliers in equilibrium under linear and nonlinear indirect network effects and evaluate the magnitude of the limitations on the platform's indirect network effects imposed by non-linearities.

We find that the number of vendors in equilibrium under nonlinear indirect network effects is significantly lower than the number of vendors under linear effects, suggesting that the growth of online platforms for local grocery food is subject to significant

[2] *Cross-side externalities* refer to the externalities exerted on one side of the platform by agents from the other side of the platform, while *same-side externalities* refer to the externalities created by the impact from one side of the platform on agents from the same side of the platform (Eisenmann *et al.*, 2006).

nonlinear indirect network economies. These non-linearities explain why, despite the fact that a number of platforms have engaged in moving fresh food directly from local producers to consumers, they have yet to dominate the provision of local foods replacing the traditional grocery channels. Therefore, although a platform business model can help connect local growers and consumers in a more efficient way, the fact that local vendors are usually highly differentiated means that the scope for any technology solution faces real economic constraints. Although the counterfactuals we propose lessen these constraints, they do not necessarily eliminate them.

The remainder of the paper proceeds as follows. In the next section, we review the literature and develop a set of testable hypotheses. We then describe our platform setting in Section 3 , as well as the data. In Section 4 , we present the model, the empirical specification, and identification strategy underlying our data analysis. In Section 5 , we present the results from both the empirical estimation exercise as well as the counterfactual simulations. We close in Section 6 with our conclusions, implications for practice and policy, and opportunities for future research.

2 Literature and Theoretical Background

Our study is relevant to the literature at the intersection of operations management, information systems, and marketing. Part of this research focuses on studying the design and growth of online platforms specialized in the sale and/or rental of durable goods in the media, electronics, automotive, home furnishings, and toys/games categories (e.g., Parker and Van Alstyne, 2005; Fraiberger and Sundararajan, 2015; Zhu and Liu, 2018) as well as in secondary markets (Dhanorkar, 2018; Richards and Hamilton, 2018; Bimpikis *et al.*, 2020). Another part focuses on the design and growth of online platform operations involving services in the lodging and passenger-transportation industries (e.g., Zervas *et al.*, 2017; Allon *et al.*, 2018).

Underlying these studies is a recognition that, for indirect network economies to emerge, a platform must ensure that participants on both of its sides obtain enough value from their involvement. Often, doing this requires the platform to generate a surplus substantial enough to be able to allocate shares of this surplus that are sufficiently large to participants to entice them to join and still leave a residual adequately large to cover the platform's costs and generate a positive rate of return (Armstrong, 2006). As a result, the higher the surplus available at the platform, the easier it is to generate large enough shares to motivate participants to join (Evans and Schmalensee, 2010, 2016).

Platforms can realize this surplus by reducing frictions involved in matching their members. The more extensive the frictions are addressed by the platforms and the greater their success in reducing them, the larger their available surplus (Evans and Schmalensee, 2016). In the case of local food platforms, these frictions are substantial. For one, there is a high degree of heterogeneity among consumers and vendors since the former often exhibit a high level of variability in taste and the latter typically specialize in producing and selling a limited range of distinct items. Moreover, exchanges between consumers and vendors are often subject to short time constraints since most consumers looking to buy these products are willing to allocate only a small fraction of their time to find them. Under these time limitations, search costs involved in matching consumers and vendors take on greater relevance. Finally, these matches are subject to information asymmetry because consumers are generally unable to assess the quality of the products offered by the vendors prior to completing their exchanges with them. Consequently, these matches are subject to a non-trivial failure rate.

In addition to the availability of a sufficiently large surplus, the emergence of indirect network economies depends on the platform attracting a sufficiently large number of members on either one of its sides to make participants on the other side to want to join in (Rochet and Tirole, 2006; Rysman, 2009). Only by scaling up both sides — the buying side and selling side — will the platform be able to generate marginal reductions in matching costs conducive to maximizing its surplus. Therefore, the platform's viability will depend on whether the surplus available to members participating on one side of the platform rises in the number of members participating on the other side and vice versa.

This phenomenon has been observed in platforms involving inventory liquidation auctions (Bimpikis *et al.*, 2020) as well as in platforms for ride-hailing services (Kabra *et al.*, 2016). However, the evidence is not uniform. For instance, a study of an online platform for short-term home rentals showed that an increase in the platform's supply side led to a lower rate of successful matchings (Li and Netessine, 2020). Moreover, evidence from an online platform matching freelance labor with local demand points to the existence of constant returns to the scale on the platform (Cullen and Farronato, 2020). In the case of a local grocery food platform, evidence to support the existence of indirect network economies is contingent on an equilibrium outcome between the vendor side and the consumer side of the platform, which has not been taken into consideration in any of the work discussed above. Yet, this equilibrium is a necessary condition to quantify indirect network economies in the growth of local grocery platforms because the growth on either side of these platforms is endogenous to the growth on the other side (Caillaud and Jullien, 2003; Hagiu, 2009). That is, consumers join in the platform based on their expectations regarding the number of available vendors and vendors participate in the platforms based on their expectations of consumer demand.

It is unclear that under this equilibrium there will be a presence of indirect network economies at a local food platform. First, the strength of indirect network effects at such platform will depend on the competitive conduct of the platform in equilibrium, since this conduct could effectively neutralize the growth potential associated with indirect network externalities (Weyl, 2010). In this setting, the platform's pricing decisions are conditioned on the equilibrium between consumer demand for local foods and supplier demand for distribution, and on the equilibrium responses by the platform's competitors. As the platform's conduct deviates from perfect competition, then its pricing decisions effectively internalize, and increasingly neutralize, any indirect network effects that may be present (Weyl, 2010). Therefore, the strength of indirect network effects depends on whether the platform's conduct deviates from perfect competition, in which case it can undermine the platform's growth potential due to indirect network externalities.

Second, increases in the demand allocated by consumers to the platform may not be sufficient to generate an increase in surplus that will draw more local vendors. This is because scalability in this type of platform is subject to added costs unique to the food industry. For one, food is unlike media, technology, advertising, or any of the other canonical platform products in that the object of the exchange (food) is difficult to store and transport, and expires relatively quickly. Furthermore, scalability in a local food platform is uniquely complex because the supply of food is seasonal and wholesale prices tend to be unusually variable (due to the relative inelasticity of both the supply of and demand for food).

Finally, scalability in a local focal platform involves users on the consumer side who care about buying products sourced from vendors located near their homes. In this environment, growth trends in demand and supply have a strong geographical coupling, making it difficult for the platform to reach a critical participation mass in equilibrium subject to indirect network economies on both of its sides (Evans and Schmalensee, 2010). Theoretical models have analyzed this phenomenon while requiring that users make their decisions to join either side of the platform in a single period (Caillaud and Jullien, 2003) or across different periods (Hagiu, 2006). These models suggest that to the extent that a platform is able to exploit these indirect network economies, it should indeed be able to scale up. Whether this is observable in a local food platform is subject to the various considerations discussed previously, in which case we could hypothesize that:

Hypothesis 1: *Indirect network economies are present in the growth of online platforms for local grocery food.*

Another consideration is that indirect network economies in a local food platform may not necessarily grow linearly in the number of members participating in a platform (Rochet and Tirole, 2003; Parker and Van Alstyne, 2005; Rochet and Tirole, 2006). There may be instances in which adding too many members to one side of a platform will impose negative externalities on the platform. Rysman (2004) and Chu and Manchanda (2016) are among the few authors who have studied empirically this phenomenon. They found that, in the case of consumer search, utility exhibits marginally decreasing returns in the variety of options consumers have available to choose from. OM scholars have studied this phenomenon but only through stylized analytical models (Halaburda *et al.*, 2017), which obviously cannot substantiate whether consumers' utility will exhibit marginally decreasing returns in the number of local food vendors participating in a platform. In theory, it is possible that an abundance of vendor choices will make it harder for consumers to find optimal food supply options. This is because when consumers are exposed to too many choices, they will engage in wasteful search and will often end up settling for suboptimal consumption decisions, including not purchasing at all (Iyengar and Lepper, 2000;

Kuksov and Villas-Boas, 2010; Arnosti *et al.*, 2021). If so, these negative *cross-side externalities* may limit the number of vendors that can participate in the platform.

Moreover, because adding vendors will increase exponentially the complexity of the network of locations where the platform will source food from, distribution costs will increase non-linearly in the number of vendors. This phenomenon is a reflection of negative *same-side externalities* among vendors in the platform and is similar to that described by Bhargava *et al.* (2013) to model analytically the costs of commercialization of products subject to network effects in demand (e.g., electric cars, video game consoles) and by Jiang and Tian (2016) and Tian and Jiang (2018) to model analytically the costs in the manufacturing and distribution of durable goods for collaborative consumption through a platform. The implication is that matching outcomes may become exceedingly more costly for the platform as the number of affiliated vendors increases because the cost of distribution incurred by the platform marginally increases in the number of vendors. As a result, these negative same-side externalities will limit the number of vendors that can participate in the platform.

From these arguments, we hypothesize below the existence of nonlinearities in indirect network economies imposed by two different types of externalities driven by the number of participating vendors in local food platforms:

Hypothesis 2a: *The growth of online platforms for local grocery food is subject to nonlinear indirect network economies imposed by negative cross-side externalities as the number of participating vendors increases.*

Hypothesis 2b: *The growth of online platforms for local grocery food is subject to nonlinear indirect network economies imposed by negative same-side externalities as the number of participating vendors increases.*

3 Methodology

Our study uses data obtained during a period of 129 weeks, starting in 2008 at the time of the platform's inception, until 2011. We selected a 129-week period after the platform's inception because most of the growth in participation on the supply side occurred within this period of operation. Since the majority of the vendors that ended up participating on the platform had already signed up by week 90, our focus on this period of analysis allowed us to ameliorate potential biases in our evaluation caused by right-censoring in the data.

3 .1 The Platform

From the start, managers at the platform focused on satisfying consumers' demand for local food because they knew that while consumers had developed a preference for these products, shopping for them across a wide variety of local vendors entailed a substantial amount of effort. Therefore, they saw an opportunity to use the platform to give consumers access to these products while significantly reducing their fixed shopping costs. The platform focused on a market in the Southeastern US where sources of local food were relatively close to consumers who would be interested in using the platform to buy from them. The market had 100,000 households, an annual population growth rate of almost 1%, and a population density of about 4,000 people per square mile. The median income per household was just over $64,000 and 25% of households had children under the age of 18 living with them. Since its launch, the platform took on the responsibility of collecting the products ordered by consumers directly from vendors and delivering these products to consumers. In turn, the vendors became responsible for picking the products consumers ordered through the platform and have them packed and ready for pick up on the morning after they received the orders. Once the platform collected all products from vendors, it assembled the orders and dispatched them for delivery to consumers in the afternoon. This ensured that consumers received their deliveries the day after they ordered through the platform.

Management also recognized that consumers preferred purchasing local groceries if the platform could meet at least some of their shopping needs for groceries from national vendors familiar to consumers. Therefore, to gain initial traction among consumers at the time it launched, the platform made available products from eight national vendors along with products from ten local vendors. The distinction between national and local vendors is a function of the vendors' locations (i.e., where their products originate) relative to the platform's market area. For example, *Eggland's Best* is considered a national vendor because its products originated from locations far removed from the market area where the platform operated. On the other hand, a vendor such as *Meadow Run Farms* is considered a local vendor since its products originated at a location in close proximity to the market area served by the platform. [3]

Soon after its launch, the platform continued expanding the number of participating local vendors while maintaining the number of national vendors essentially unchanged. While the number of local vendors increased over time until leveling at around 65 by week 90, the national vendor base remained essentially unchanged during our data collection period. As the number of local vendors continued increasing during this period, management kept updating a list of these vendors and prominently displayed it on the front page of the platform's online site.

[3] This characterization is based on that used by the US Congress in the 2008 Farm Act, which considers food transported less than 400 miles, or that is sold within the state where it is grown, to be locally sourced (Darby *et al.*, 2008; Martinez, 2010).

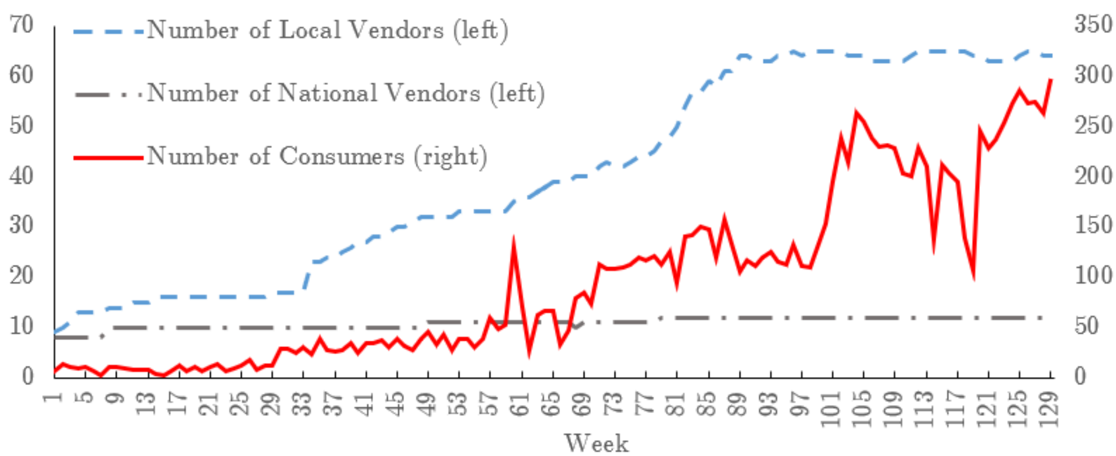
3 .2 Data description

Our data consist of $34,327$ consumer-level transactions on the platform from September of 2008 until March of 2011. For each transaction, the data included the consumer's identification number as well as a description, count, and price for each product purchased, along with the cost paid by the platform to the vendors for each product. The transaction data also included the date of the purchase, the dollar amount in coupons the consumer used to make the purchase, and whether the consumer was paying for the delivery of the order through an optional delivery subscription program available through the platform. Furthermore, the transaction data included detailed information on the vendor selling each item. This information was also available to consumers on the platform's website. For each item, the website displayed the name of the vendor and indicated whether the vendor was local to the market. For instance, collard greens sourced from a local vendor (Meadow Run Farms) carried a label that read "Meadow Run Farms Produce-Local No-Spray Collard Greens". The use of labels advertising the name of the vendors and specifying whether the vendors were local to the market not only allowed the platform to inform consumers explicitly of the products' provenance, but also gave consumers the ability to filter their search for products available on the platform to include only those supplied by local vendors.

From these data, we built Figure 1 to illustrate the change in the composition of the platform's vendor base: While the number of local vendors increased steadily over time until leveling at week 90, the national vendor base changed only slightly in weeks 8, 49, and 79. To control for the changes in the national vendor base, we used the first period (week 1 to week 7) as the base level and assigned a dummy to the second (week 8 to 48), the third (week 49 to 78), and the last period (week 79 to 129), respectively.

In a two-sided platform like the one in our study, indirect network economies would create a virtuous cycle in which an increase in consumer demand at the platform would facilitate an increase in the number of suppliers participating on the platform and vice versa. Figure 1 provides visual evidence that both the number of consumers and the number of vendors on the platform are positively related. Although this is an indication that indirect network economies may exist in the platform, it fails to properly measure consumer demand as well as to account for potential endogeneity between consumer demand and the number of vendors at the platform as well as other confounding effects that may contribute to explain this relationship. In the next section, we expand on the econometric model we used to account for these issues.

Figure 1: Size of vendor and consumer base by week



4 Econometric Model

We use a structural model in which we estimate consumer *weekly* demand at the platform in the first stage, and the marginal effect of an increase in consumer demand on the number of local vendors in the second stage. By endogenizing both the consumer demand and the supply of local groceries on the platform, we estimate the strength of participation on each side of the platform and identify the presence of indirect network economies along with the number of vendors at the observed equilibrium.

To model consumer weekly demand at the platform, we follow the same approach as Nair *et al.* (2004) and Richards and Hamilton (2018) and use a simulated maximum likelihood estimation of the joint probability of each consumer shopping on the platform on a weekly basis (in Section 4 .1) and of the number of items purchased by the consumer (i.e., the basket size) given the consumer's choice of shopping on the platform during a particular week (in Section 4 .2). The probability of purchasing is driven by considerations of need and state dependence, while basket-size demand is driven more by volume-demand. Both the probability of choosing the platform and the basket size depend on the number of local vendors participating on the platform and the price of items sold through the platform. However, because these determinants may correlate with unobserved demand shocks, we use a control function approach (which we detail in Section 4 .3) to account for their potential endogeneity. [4]

The modeling approach reflects the following sequential decision process of consumers. In a given week, a consumer first decides whether or not to make a purchase on the platform. As consumers seek to match their specific preferences on the platform, the probability of finding an ideal match increases at a decreasing rate as there are diminishing marginal returns to variety (Mehta *et al.*, 2003; Draganska and Jain, 2005; Richards *et al.*, 2015; Chu and Manchanda, 2016). Therefore, we expect to see a

[4]We chose a control function approach over a two-stage least squares (2SLS) approach because the former constitutes a superior approach to obtaining an instrumental variable estimator (Chan *et al.*, 2020). In a standard case where an endogenous variable appears linearly, the control function will lead to the usual 2SLS estimator (Wooldridge, 2015). However, in the case of a nonlinear model (as is the case in our paper), the control function approach is at an advantage relative to the 2SLS approach because the key exclusion assumption for a valid instrumental variable is not restricted to its zero correlation with the error term but to its conditional independence. This conditional independence allows the control function to be included in the regression estimation in a flexible manner rather than in a pre-specified function form, such as that based on the linear projection of the endogenous variable on the instrumental variable in the 2SLS approach (Petrin and Train, 2010). Therefore, compared to 2SLS, the control function is a superior approach to tackle endogeneity in the estimation of our paper's nonlinear equations on the demand side.

nonlinear impact of variety on the probability of choosing whether to purchase. The consumer then decides how many units to buy in a basket. Conditional on having made the decision to purchase, the second-stage basket-size decision reflects more "traditional" volumetric concerns (i.e., the demand curve from economic principles) and not the diminishing marginal probability of achieving a preference match (Bucklin *et al.*, 1998; Ailawadi *et al.*, 2007). In this sense, there is no reason to expect the same nonlinear relationship between variety and quantity-purchased as we expect at the purchase-incidence stage.

By aggregating the product of purchase incidence and the expected number of items purchased per basket across all consumers per week, we obtain a predictive model of weekly platform demand, which we then use to estimate a model of equilibrium product provision at the platform per week (in Section 4 .4). The goal of this model is to examine how local vendors respond to increases in consumer demand at the platform. We estimate this model using a generalized method of moments (GMM) with instruments to address the potential endogeneity of the surplus and the number of participating local vendors. This approach is consistent with that used by Villas-Boas and Zhao (2005) and Richards and Hamilton (2015). We expand in Section 4 .4 on the use of these instruments as well.

4 .1 Purchase Incidence Model

We model the number of orders on the platform each week, t, as the product of the total number of consumers visiting the platform and the probability that each consumer purchases on the platform in that week. This probability depends on platform attributes that include one of our primary variables of interest, the number of local vendors participating in the platform, N_t. In week t, consumers choose to either purchase on the platform or purchase elsewhere (or not at all). In the literature, the latter correspond to an outside option with a utility of zero (Berry *et al.*, 1995; Villas-Boas and Zhao, 2005).

When consumer h visits the platform and purchases products from the platform in week t, she obtains a utility given by:

$$U_{ht} = \gamma Z_{ht} - \alpha p_t + \beta X_t + f(N_t) + \varepsilon_{ht}, \tag{1}$$

where p_t denotes the average price of products available on the platform and ε_{ht} is an iid random error term. The utility function contains a vector X_t of time-related variables. The vector contains a variable counting the number of weeks since the platform's inception (*WEEK*) in order to control for longitudinal growth in demand and a dummy variable (*BREAK*) indicating whether a given week, t, occurs during times of the year when consumers typically leave for vacation (between Memorial Day and Labor Day and between Christmas Day and New Years' Day). [5] Since consumers' value of time is lower during these periods, we expect to observe a lower U_{ht} during these weeks. The vector also contains the dummy variables that identify the periods in our analysis during which the size of the national vendor base remained constant. We expect to observe a different value of U_{ht} in each of these periods.

The utility function also includes a vector, Z_{ht}, of consumer need-based variables, or variables that proxy the likelihood that in-home inventories of certain items may be running low. The first of these variables corresponds to a lagged quantity (LQ) measure of the number of items in the consumer's most recent purchase on the platform. We expect that a higher LQ will have a negative effect on the utility of purchasing at the platform during a given week. The vector also includes a measure of inter-purchase time (IPT), in weeks, corresponding to the interval between the most recent purchase on the platform by the consumer and the current week, t. IPT reflects the level of

[5] We include the time trend in the purchase incidence model to control for state-dependence such as inertia, learning or habituation that our other measures do not capture.

loyalty that consumers have formed towards the platform and therefore as it increases, the utility of purchasing is likely to decrease. Furthermore, the vector includes a consumption rate (CR) measure equal to the total number of items purchased by the consumer on the platform during our observation period divided by the number of weeks the consumer stayed active on the platform. Thus, we implicitly assume that consumers do not know about the platform before their first transaction and choose not to shop at the platform again after their last transaction. We expect a consumer with a high CR to be more likely to regard the platform as his or her primary source for groceries and this will imply a higher utility of purchase (Bell *et al.*, 1998; Briesch *et al.*, 2009).

The Z_{ht} vector also includes a measure that captures the type of delivery payment plan (DPP) that consumers chose in order to have their platform purchases delivered to them. Although the platform did not offer rush deliveries during our period of analysis, it did provide consumers with the option to pay in advance a membership fee to be eligible for an unlimited number of deliveries on a monthly basis. Consumers who chose not to use the membership plan had to pay for delivery every time they ordered on the platform. The measure we included in the Z_{ht} vector differentiates between consumers who chose the membership plan and consumers who chose to pay for delivery every time they purchased on the platform. We expect to observe a higher utility for consumers who chose the membership plan than for consumers who did not. The former are likely to have a higher utility than the latter.

Recall from **Hypothesis 2a** that vendor variety available on the platform may have a nonlinear effect on consumer utility. We capture this effect by following Draganska and Jain (2005) and Richards and Hamilton (2015) and use a quadratic form $f(N_t) = \theta_{1h} N_t + \theta_{2h} {N_t}^2$ to define the attraction for consumers to the platform as a nonlinear function of the number of local vendors on the platform. The derivation of this subutility function can be found in both Draganska and Jain (2005) and Richards and Hamilton (2015). Furthermore, it accounts for the possibility of unobserved heterogeneity among consumers' preferences for vendor variety on the platform. Therefore, it allows the consumer-level parameters capturing the effects of N_t and N_t^2 in Equation (1) to vary across consumers such that:

$$\theta_{1h} = \theta_1 + \sigma_1 \nu_1 \qquad \nu_1 \sim \ N(0,1) \tag{2}$$

$$\theta_{2h} = \theta_2 + \sigma_2 \nu_2 \qquad \nu_2 \sim \ N(0,1), \tag{3}$$

where θ_{1h} and θ_{2h} are normally distributed with mean of θ_1 and θ_2. Consumer heterogeneity in the preference for vendor variety is distributed according to a normal distribution with zero mean and standard deviation of σ_1 and σ_2, which we introduce to the model explicitly by standardizing two random variables, ν_1 and ν_2, such that $\nu_1 \sim N(0,1)$ and $\nu_2 \sim N(0,1)$. Hence, the parameters σ_1 and σ_2 capture the heterogeneous consumers' preferences for vendor variety.

Each week, a consumer decides whether to buy from the platform or not. We assume the unobserved consumer tastes in Equation (1) follow a Type I extreme value distribution, so each consumer's decision is described by a logit probability distribution. Therefore, the probability of purchase incidence for consumer h in week t is given by:

$$\Pr(inc) = \frac{\exp\left(\gamma Z_{ht} - \alpha p_t + \beta X_t + \theta_{1h} N_t + \theta_{2h} {N_t}^2\right)}{1 + \exp\left(\gamma Z_{ht} - \alpha p_t + \beta X_t + \theta_{1h} N_t + \theta_{2h} {N_t}^2\right)}. \tag{4}$$

Because we account for unobserved consumer heterogeneity in the preferences for vendor variety in Equations (2) and (3), the logit model in Equation (4) does not have a closed-form solution. Therefore, we integrate over the distributions of θ_{1h} and θ_{2h} in Equation (4) by taking 100 Halton random draws from their distribution (Train, 2009). Then, from the first and second order effects for N_t estimated in this solution, we can determine how consumers' likelihood of purchasing on the platform changes with the number of local vendors.

4 .2 Basket Size Model

The dependent variable in the basket size model corresponds to the number of items that a consumer buys. We model this quantity decision as zero-truncated Negative Binomial, following the approach used in Greene (2003). This decision is conditional on the consumer's purchasing on the platform. Therefore, given that a consumer h, makes a purchase on the platform in week t, the probability of purchasing $q_{ht} = 1, 2, \ldots, n$ units is written as:

$$P(Q_{ht} = q_{ht}|Q_{ht} > 0) = \frac{\exp(-\lambda_{ht})(\lambda_{ht})^{q_{ht}}}{(1 - \exp(-\lambda_{ht}))\, q_{ht}!}, \tag{5}$$

where $\lambda_{ht} = \exp(\psi_{h_0} + \psi_p p_t + \psi_n N_t + \psi_I I_{ht})$ is the purchase rate of consumer h on the platform in week t. Conditional on the consumer having chosen to make a purchase, basket size is likely to decrease in p_t. Furthermore, following Richards and Hamilton (2018), we assume that basket size is likely to increase in N_t and vary with the vector I_{ht}. This vector includes the consumer need-based variables introduced in Section 4 .1, as well as a variable that accounts for consumers' use of promotional programs at the platform, as reflected in the dollar amount consumers used in coupons *(COUPON)* provided by the platform as part of their purchases in week t. Finally, ψ is a vector of parameters to be estimated and $\exp(.)$ ensures that the purchase rate will be non-negative. As in the purchase incidence model, we allow for unobserved heterogeneity over consumers by allowing the intercept term to be normally distributed such that $\psi_{h_0} = \psi_0 + \psi_1 v_3, v_3 \sim N(0, 1)$, where v_3 is an independent and standardized random variable.

We considered using a Poisson distribution to model consumers' basket size decisions. However, this distribution proved to be too restrictive for this purpose because it assumes that the mean of the estimated basket size is equal to its variance. In our setting, this assumption does not hold due to over dispersion among basket size observations. Specifically, in a Poisson model the relationship between mean and variance is: $Var\left[q_{ht}|p,N,I\right] = E\left[q_{ht}|p,N,I\right]$, whereas in a negative binomial model the relationship is: $Var\left[q_{ht}|p,N,I\right] = E\left[q_{ht}|p,N,I\right] + \alpha E\left[q_{ht}|p,N,I\right]^2$ which indicates that the variance is greater than the mean when α is greater than zero. Following Cameron and Trivedi (1990), we performed a t-test to verify whether the mean of the estimated basket size is equal to its variance. The results from this test led us to reject the null hypothesis, and conclude that overdispersion is indeed a feature of our data. Because the Negative-Binomial-P's density function in Equation (6) is unrestricted, we can estimate its parameters, α (overdispersion) and Q (form), from the data:

$$P(Q_{ht} = q_{ht}|p,N,I) = \frac{\Gamma\left(T{\lambda_{ht}}^{Q} + q_{ht}\right)}{\Gamma\left(T{\lambda_{ht}}^{Q}\right)\Gamma\left(q_{ht}+1\right)} \left(\frac{T{\lambda_{ht}}^{Q}}{T{\lambda_{ht}}^{Q} + \lambda_{ht}}\right)^{T{\lambda_{ht}}^{Q}} \left(\frac{\lambda_{ht}}{T{\lambda_{ht}}^{Q} + \lambda_{ht}}\right)^{q_{ht}}. \tag{6}$$

In this model, T is an estimate of $\frac{1}{\alpha}$. If α approaches 0, the Negative Binomial-P will collapse to Poisson. If $Q = 0$ or $Q = 1$, the functional form becomes a *Negbin2* or a *Negbin1*, respectively.

4 .3 Demand Estimation and Identification

We estimate the parameters in the purchase incidence model and basket size model jointly through a simulated maximum likelihood estimation procedure (Train 2009). This approach is based on the following simulated log-likelihood (SLL) function:

$$SLL = \sum_{h=1}^{H} ln \left\{ \frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T} [Pr_{ht}(inc) P(Q_{ht} = q_{ht} | Q_{ht} > 0)]^{y_{ht}} \left[1 - Pr_{ht}(inc)\right]^{(1-y_{ht})} \right\}, \tag{7}$$

where y_{ht} equals 1 if consumer h chooses to buy on the platform in week t or 0 otherwise. As we mentioned above, we allow individual-specific parameters θ_1 and θ_2 in the utility function and ψ_1 in the basket size function to vary randomly across consumers and, thus, we use Halton draws from the population distribution for $r = 1, 2, \ldots, R$. This procedure has been used in the literature to improve estimation efficiency (Train 2009). We observed little difference in the parameter estimates after more than 100 draws.

In structural models such as this, identification is aided by non-linearity in the functional forms for each part, but is not guaranteed due to the presence of clear endogeneity. Specifically, the number of participating vendors and the price of groceries in the purchase incidence and basket size models may be correlated with unobserved shocks affecting consumers' demand, violating the conditions necessary for their assumed exogeneity. Therefore, to address these predictors' potential endogeneity, we used a control function approach (Petrin and Train, 2010) with two instrumental variables (IVs). We instrument price with a well-accepted cost shifter: the wholesale cost incurred for every product sold on the platform. It is well known that wholesale cost correlates strongly with price. Moreover, because the platform was an extremely small player relative to other firms, particularly grocery retailers, that sold similar products to consumers, its wholesale costs are highly likely to be mean-independent of demand shocks. Therefore, we retained the residual of this regression as a price control, which is unobserved and not explained by the observed choice characteristics and the IV, and included it in Equations (4) and (6).

We carefully evaluate the validity of wholesale cost as an IV to ensure it satisfies both the inclusion and exclusion restrictions. First, the inclusion restriction ensures that the IV is correlated with the explanatory endogenous variable. To that end, we test for the strength of our instrument using an F-test from the first-stage instrumental-variables regression (Staiger and Stock, 1997; Stock *et al.*, 2005). As shown in the appendix, the F-statistic (29,887.35 and 29,885.94) and R^2 values (0.87 and 0.87) obtained in the first-stage IV regression of price on wholesale cost and the exogenous explanatory variables in the purchase incidence and basket size models, suggest that wholesale cost is not a weak instrument for price. Second, the exclusion restriction ensures that the IV is not correlated with the error term in the estimation equation. Exclusion restrictions cannot be validated through formal statistical tests, as the error term is unobserved, but must be supported by theoretical or empirical evidence from previous research (Ho *et al.*, 2017). In this regard, wholesale costs have been commonly used as instruments in the empirical literature studying grocery sales to consumers (e.g., Chintagunta, 2002; Sriram and Kalwani, 2007). The exclusion restriction for using wholesale cost also has theoretical support not only because the platform was an extremely small player relative to its competitors, but also because the food supplier sector tends to be relatively competitive, which suggests that wholesale price more closely reflects the marginal cost of production (Ailawadi *et al.*, 2010). Marginal cost, in turn, is clearly independent of demand shocks (Villas-Boas, 2007; Nakamura and Zerom, 2010). [6]

[6] Nakamura and Zerom (2010) cited a few early empirical studies that assume manufacturer-Stackelberg in vertical strategic interactions, which implies greater manufacturer power. But it also cited more recent research showing how the dominance of retailers' strategies selling differentiated products such as private label products has shifted bargaining power to retailers. Retailer concentration, private label, and hard-discounter entry mean that wholesale prices have become highly competitive over the past couple of decades.

The instrument we chose for the number of local vendors participating on the platform consists of weekly volume of food items shipped by local vendors from the area where the platform's market is located to other parts of the country. This instrument exploits inherent variations in the availability of products for sale at the platform from local vendors as a result of their preset commitments in traditional channels, as well as weather conditions and other biological factors that are largely beyond the purview of managers at the platform. We further validate why the volume of food items shipped by local vendors is a good instrument. First, it reflects the supply of local foods in a much larger market than the one that generates the data in our study (i.e., the entire US). In this sense, our instrument reflects a similar logic to Hausman *et al.* (1994) and DellaVigna and Gentzkow (2019), who argue that because retail prices are set on a national level, they are exogenous to demand in any particular market. In our case, supplies are determined by national considerations, so they must be mean independent of the weekly demand shocks of individual consumers on the platform of a small market. Second, because shipment volumes reflect local vendors' preset commitments in traditional channels and planting decisions for most products are made many weeks, sometimes months, in advance of the shipment data, they cannot be contemporaneously correlated with shocks in demand for a specific item (Ahumada and Villalobos, 2009). Thus, the shipment volumes in this instrument are exogenous to weekly demand shocks at the platform but will correlate negatively with the weekly number of vendors participating on the platform.

Our measure of weekly shipments consists of weekly volumes (in $10,000$ lb. units) recorded by the US Department of Agriculture's (USDA) Agricultural Marketing Service for a variety of produce shipments sent from the area where the market served by the platform during our period of analysis is located (https://www.ams.usda.gov/). We then used these data to create an index for each of the 129 weeks in our analysis corresponding to the straight sum of all weekly shipment volumes for the top 15 items sold on the platform during our period of analysis. Specifically, to create the index, we added the weekly produce shipment volumes (in $10,000$ lb. units) matching the top selling items on the platform. Because of the large amount of products offered through the platform, matching each product with every item recorded in the USDA's shipment volume reports becomes intractable. Therefore, the shipment volumes of the top selling items constitute a more reliable reflection of the overall outbound volumes from the market. The F-statistic (31,558.49 and 31,186.52) and R^2 values (0.88 and 0.88) we obtained when we ran the first-stage IV regression using this index in both the purchasing incidence and basket size models provide confirmation that the index is not a weak instrument (Staiger and Stock, 1997). Please refer to the appendix.

4 .4 Equilibrium of Supply Provision

In this stage of the analysis, we model the number of local vendors and their product prices on the platform. Our modeling approach is consistent with Weyl (2010) as it emphasizes the importance of estimating indirect network effects, conditional on equilibrium responses from firms that co-exist in a competitive industry. To that end, we estimate first-order conditions derived from profit-maximizing behavior by the platform, assuming oligopolistic competition. Although our first-order conditions are derived from a static setting, they capture outcomes of coordinating Nash equilibria (Cachon, 2003). Implicitly, static Nash equilibria assume that the equilibrium prices and number of local vendors reflect rational expectations among platform stakeholders. That is, the dynamics are subsumed in their correct expectations as to how the game is going to evolve if played out repeatedly, over time. Recent analytical studies that have applied the static equilibrium model to derive indirect network economics in two-sided markets indicate that this can be generalized to a dynamic setting (Halaburda *et al.*, 2017; Benjaafar *et al.*, 2020).

When the platform launched in 2008, it became one of the first online grocery platforms operating in the US. Therefore, most of its competition for a share of the local foods' market it served came not from other platforms but from traditional grocery retailers, specialty stores, and farmers' markets. As such, we model the platform's choice of price and the number of local vendors as a Bertrand-Nash equilibrium in both variables. [7] Following the empirical game described by Richards and Hamilton (2015, 2018), management chooses each week the number of local vendors that will participate in the platform and the prices for the products sold through its website. [8]

Focusing only on our single platform, the profit equation is given by:

$$\pi_t = E\left[Q_t\right]\left(p_t - c_t\right) - g\left(N_t\right), \tag{8}$$

where $E[Q_t]$ is the expected platform demand aggregated from the demand estimation. $E[Q_t]$ corresponds to the product of the market size (measured as the population of households in the market), the average probability of purchasing, and the expected item count in an individual basket. p_t is the average price on the platform in week t and c_t denotes the marginal retailing costs for the platform. We estimate c_t from the data as $c_t = kW_t + \zeta_t$. In this function, W_t is a vector of cost shifters comprising the average wholesale cost incurred by the platform when paying to the vendors for the items sold every week as well as retail hourly wages and utilities' costs collected from

[7] In the model, we implicitly capture the extent of competition between the focal platform and its competitors as absorbed by the outside option in the demand model.

[8] This is consistent with common practice among local grocery platforms, including the one in our study. Based on this practice, the platform sets prices via a markup over wholesale prices charged by local vendors, and this markup determines the selling prices charged to consumers. However, the price at the platform will affect not only consumers' probability of purchasing and the size of their baskets but also the surplus to be allocated to local vendors and these vendors' decision to participate on the platform. As a result, the pricing and the number of participating vendors at the platform are also an outcome of the model's equilibrium, not just the result of decisions by the platform.

the US Bureau of Labor Statistics during our period of analysis and ζ_t is a random supply shock.

Finally, $g(N_t)$ are the costs the platform incurs matching consumers and vendors. These include distribution costs as well as subsidies and other incentives paid to vendors for every transaction with consumers. To model these costs, we use a quadratic function defined as $g(N_t) = \delta_0 N_t + \frac{1}{2}\delta_1 N_t^2$ to account for the possibility of $g(N_t)$ being a convex function, per **Hypothesis 2b**.

The platform's economic surplus is determined by both p_t and N_t, so the first order conditions (FOCs) in prices and number of participating local vendors become:

$$\frac{\partial \pi_t}{\partial p_t} = E[Q_t] + \frac{\partial E[Q_t]}{\partial p_t}(p_t - c_t) \tag{9}$$

$$\frac{\partial \pi_t}{\partial N_t} = \frac{\partial E[Q_t]}{\partial N_t}(p_t - c_t) - \frac{\partial g_t}{\partial N_t}. \tag{10}$$

Solving for price and number of local vendors and dropping the time subscripts for clarity gives:

$$p = c - \varphi(\Delta_p)^{-1} E[Q] \tag{11}$$

$$N = -\frac{1}{\delta_1}\Delta_N(\Delta_p)^{-1}E[Q] - \frac{\delta_0}{\delta_1}, \tag{12}$$

where Δ_p and Δ_N correspond to the expected-quantity-derivatives with respect to the average platform retail price and the number of participating local vendors, respectively. While Δ_P will be negative, Δ_N will be positive. In turn, φ is a parameter that captures the platform's competitive conduct form (Besanko *et al.*, 1998; Richards and Hamilton, 2006). If the estimate of φ converges to 1, the platform's conduct is consistent with Bertrand-Nash rivalry. On the contrary, if the estimate of φ converges to 0, the platform's conduct approximates perfect competition. In the latter case, none of the available indirect network effects are absorbed by the platform (Weyl, 2010).

As indicated by Equation (12), the reduced form of the size of the platform's local vendor base is a function of the product of the expected local food quantity that consumers demand at the platform, $E[Q]$, and the expected-quantity derivatives with respect to retail price, Δ_p, from Equation (11), and with respect to the number of local vendors on the platform, Δ_N. Because, in equilibrium, the marginal cost of expanding the size of the local vendor base equals the marginal surplus (corresponding to the product of $E[Q]$, Δ_p, and Δ_N in Equation (12)), a positive estimate of $\frac{1}{\delta_1}$ in Equation (12) would imply that adding an additional local vendor increases the platform's surplus available to the vendors and that expanding the size of the local vendor base is beneficial. This result would provide evidence to support **Hypothesis 1**. Moreover, based on the estimate of $\frac{1}{\delta_1}$, we will also be able to simulate the number of vendors at the platform's observed equilibrium. To determine value for the latter, we first estimate $\frac{1}{\delta_1}$ along with the other parameters in Equations (11) and (12) and then use these estimates to solve simultaneously for the number of local vendors in Equation (12) and the price in Equation (11).

Note that both Δ_p and Δ_N are likely to be endogenous since they may correlate with unobservable supply shocks that may influence decisions by management at the platform regarding pricing and the composition of participating local vendors, respectively. Therefore, we chose lagged values of Δ_p and Δ_N as IVs for these variables, since they reflect demand-side shocks but are also mean independent of the residuals of the platform's prices and supplier-base size. As shown in the appendix, the F-statistic (183.64 and 214.59) and R^2 values (0.88 and 0.77) for the equations based on these IVs suggest that none of them can be characterized as weak. In the next section, we present the results we obtained from estimations using both IV and non-IV models to demonstrate the endogeneity biases in our structural system.

5 Results

We present the results we obtained from the structural model of demand and supply provision in three parts. First, we will review the summary statistics from the data for the variables of interest in the model (in Section 5 .1). We will then present the results from the demand stage involving purchase incidence and basket size (in Section 5 .2). These results provide the basis to evaluate consumers' demand on the platform as a function of the number of participating vendors. Finally, we will present the results from the supply stage (in Section 5 .3). From these results, we will be able to test **Hypothesis 1** by quantifying the marginal effect on the platform's surplus from expanding the number of participating local vendors while controlling for the simultaneous effects of platform prices and demand on the size of the platform's local vendor base. Moreover, we will be able to estimate the number of participating vendors in the observed equilibrium and whether the indirect network effects in the platform are nonlinear in the number of participating vendors. The results from this analysis will provide the basis to test **Hypothesis 2a** and **Hypothesis 2b** through counterfactual simulations of variations in cross-side and same-side externalities (in Section 5 .4)

5 .1 Summary Statistics

Table 1 presents the variable descriptions and summary statistics from the data. One of our primary variables of interest, the number of local vendors, varies from a low of 10 to a high of 65 over the period of analysis. In contrast, the number of national vendors maintains a minimal variation, ranging from 8 to 12 during the sample period. The high level of variation in the number of local vendors relative to that in the number of national vendors suggests that the effect of the number of local vendors is well identified.

Table 1: Summary statistics and variable descriptions

Variable	Description	Mean	Std. Dev.	Min.	Max.
Number of Local Vendors	The number of local vendors available on the platform per week	40.58	19.75	10	65
Number of national vendors	The number of national vendors available on the platform per week	10.23	0.94	8	12
Average price	Average price of all items offered on the platform per week	3.64	0.33	2.91	4.35
Choice	The binary choice to purchase from the platform (1= Yes, 0= No)	0.29	0.46	0	1
Basket Size	The number of items purchased per basket per order, per week	27.17	19.50	1	239.00
Wholesale Cost	Average wholesale cost of all items offered on the platform per week	3.22	0.29	2.52	4.03
Coupon Amount *(COUPON)*	The dollar amount in coupons used per order per order, per week	0.59	2.67	0.00	102.78
Consumption Rate *(CR)*	The ratio of total purchased quantity to the number of active weeks per order, per week	8.9	10.21	0.08	104.28
Inter-Purchase Time *(IPT)*	The number of weeks since last purchase per order, per week	7.84	11.68	1	97.00
Vacation Week *(BREAK)*	The week includes times of the year when consumers typically leave for vacation (1= Yes, 0= No)	0.30	0.46	0	1
Delivery Payment Plan *(DPP)*	Order placed by consumers with delivery subscription membership (1= Yes, 0= No) per order, per week	0.04	0.20	0.00	1

Furthermore, the statistics in Table 1 suggest that there is enough variation in the data to identify the key demand parameters at the platform. Table 1 shows that consumers purchase a wide range of basket sizes since the number of items purchased (i.e., the number of SKUs multiplied by the number of units per SKU) per basket varies from 1 to 239 and has an average of 27.17. Table 1 also shows that consumers have an average consumption rate of 8.9 items per week which varies from a low of 0.08 items to a high of 104.28 items per week. The large variation in basket sizes and weekly consumption rates points to a possible presence of consumer heterogeneity where some consumers use the platform to fulfill a substantial proportion of their grocery needs while others shop at the platform occasionally to top off their traditional grocery purchases. This observation is also reflected in the statistics for the binary variable measuring consumers' choices to buy at the platform. On average, consumers buy at the platform 30% of the time, assuming a weekly shopping cycle. Moreover, according to the standard deviation for this variable, this frequency varies substantially across consumers.

Finally, the averages for product prices (\$3.64) and costs (\$3.22) in Table 1 give a sense of these variables' orders of magnitude. Moreover, the difference between these two averages gives an indication of how large the markups are on the platform. Table 1 also offers insights into consumers' average use of delivery membership plans and coupons at the platform. Almost, 5% of consumer purchases per week use the delivery membership plan. Moreover, on average, consumers use \$0.59 dollars in coupons per purchase, per week.

5 .2 Structural Model: Demand Stage

We first used simulated maximum likelihood to jointly estimate purchase incidence and basket size in the demand stage of the structural model. As explained in Sections 4 .1 and 4 .2, we used a logit approach to model consumers' decisions to purchase at the platform and a Negative Binomial approach to model consumers' basket-size decisions once consumers have decided to purchase at the platform. Table 2 presents the results from our analyses across three different model specifications, labeled in Table 2 as Model 1, Model 2, and Model 3, respectively. Table 2 also lists the labels we use to identify the predictors in our discussion below.

Contrasting the Specifications in Models 1–3

The results from the first specification (in Model 1) do not control for consumer heterogeneity or account for any sources of endogeneity of either platform price or the number of participating local vendors. The results from the second specification (in Model 2) account for endogeneity through the control functions for price and the number of participating local vendors in Section 4 .3. Because the estimated effects for the predictors in the basket size model are conditional on the effects estimated in the purchase incidence model, we include the controls for price and number of local

Table 2: Demand estimation

	Model 1		Model 2		Model 3	
Purchase Incidence	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Constant	-1.5340***	0.3737	-1.4304***	0.3966	-1.4301***	0.3966
Average Price	-0.2323***	0.0702	-0.3564***	0.0799	-0.3564***	0.0799
Number of Local Vendors θ_1	0.0718***	0.0143	0.1104***	0.0163	0.1104***	0.0163
Number of Local Vendors σ_1					-0.0003	0.0002
Squared Number of Local Vendors θ_2	-0.0011***	0.0002	-0.0010***	0.0002	-0.0010***	0.0002
Squared Number of Local Vendors σ_2					-1.94E-06	3.75E-06
Lagged Quantity *(LQ)*	-0.0027***	0.0009	-0.0027***	0.0009	-0.0027***	0.0009
Inter-Purchase Time *(IPT)*	-0.0681***	0.0018	-0.0685***	0.0018	-0.0685***	0.0018
Consumption Rate *(CR)*	0.0707***	0.0018	0.0740***	0.0019	0.0740***	0.0019
Delivery Payment Plan *(DPP)*	0.4584***	0.0599	0.3963***	0.0615	0.3964***	0.0615
Vacation Week *(BREAK)*	-0.2083***	0.0382	-0.2006***	0.0398	-0.2003***	0.0398
Week Dummy 1	-0.5576**	0.2535	-0.6241**	0.2528	-0.6239**	0.2528
Week Dummy 2	-1.0411***	0.2977	-0.9571***	0.2971	-0.9566***	0.2971
Week Dummy 3	-0.5502*	0.3177	-0.6354**	0.3168	-0.6352**	0.3168
Week (*Trend*)	0.0152***	0.0014	-0.0078	0.0052	-0.0078	0.0052
Price Control			1.0330***	0.2114	1.0337***	0.2115
Number of Local Vendors Control			-0.0514***	0.0105	-0.0514***	0.0105
Basket Size						
Constant	3.3623***	0.1039	2.9006***	0.1191	2.9004***	0.1195
Average Price	-0.1969***	0.0287	-0.0876***	0.0324	-0.0879***	0.0325
Number of Local Vendors	0.0079***	0.0005	0.0087***	0.0005	0.0088***	0.0005
Lagged Quantity *(LQ)*	0.0090***	0.0002	0.0090***	0.0002	0.0090***	0.0002
Inter-Purchase Time *(IPT)*	0.0007	0.0010	0.0006	0.0010	0.0005	0.0010
Delivery Payment Plan *(DPP)*	0.2612***	0.0151	0.2682***	0.0148	0.2680***	0.0149
Coupon Amount *(COUPON)*	0.0170***	0.0005	0.0172***	0.0005	0.0172***	0.0005
Vacation Week *(BREAK)*	-0.0263**	0.0133	-0.0233*	0.0132	-0.0235*	0.0133
Price Control			-0.1958**	0.0953	-0.1967**	0.0954
Number of Local Vendors Control			-0.0137***	0.0011	-0.0137***	0.0011
ψ_1					0.0021	0.0056
T	0.0456***	0.0074	0.0410***	0.0065	0.0421***	0.0067
Q	1.2576***	0.0491	1.2962***	0.0485	1.2884***	0.0484
N	34327		34327		34327	
LL	-58711.71		-58591.20		-58590.31	
AIC/N	3.422		3.415		3.415	

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10%.
Model 2: Controls for endogeneity of price and number of local vendors.
Model 3: Controls for both endogeneity and consumer heterogeneity.

vendors in both the purchase incidence and basket size models to remove any biases caused by these predictors' endogeneity. Finally, the third specification (Model 3) adds random consumer level parameters to the second specification in order to account for consumer heterogeneity.

With a couple of exceptions, a comparison of the results across all model specifications reveals a high degree of consistency in the signs, magnitude, and statistical significance among the parameter estimates corresponding to the predictors' effects. These exceptions are the parameter estimates of the effects of local vendor base size and price obtained from the first specification relative to those obtained in the other two specifications. The discrepancies in these estimates are likely the result of unaccounted biases caused by endogeneity in the first specification.

We find that the coefficients corresponding to the number of local vendors when we control for endogeneity in the second and third specifications of the purchase incidence ($\theta_1 = 0.1104, p < 0.01$ in Model 2 and Model 3) and the basket size models ($0.0087, p < 0.01$ in Model 2 and $0.0088, p < 0.01$ in Model 3) are of the expected sign and statistically significant. Moreover, these coefficients are higher than those obtained without control functions in the first specification of the purchase incidence ($0.0718, p < 0.01$) and the basket size models ($0.0079, p < 0.01$). This indicates a downward bias in the estimated coefficient for the effect of local vendors on consumers' demand when no endogeneity is taken into account.

Similarly, although the coefficients for price obtained from all three specifications of the purchase incidence and the basket size models are of the expected sign and statistically significant, their values in the second and third specifications differ markedly from those of the coefficients in the first specification. In the purchase incidence model, the coefficient in the first specification ($-0.2323,\ p < 0.01$) has a value that exceeds those obtained for this coefficient in the other two specifications

(-0.3564, $p < 0.01$ in Model 2 and Model 3). This suggests that the estimate of the effect by price on purchase probabilities is upward biased without correcting for price endogeneity at the platform. In the basket size model, on the other hand, the coefficient for price in the first specification (-0.1969, $p < 0.01$) is lower than those in the other two specifications (-0.0876, $p < 0.01$ in Model 2 and -0.0879, $p < 0.01$ in Model 3), suggesting a downward bias in the estimate of the effect by price on basket size without correcting for price endogeneity. [9]

Based on these considerations, the estimates in the second and third specifications are better suited than those in the first specification to evaluate the purchase incidence and basket size models. Nevertheless, when we compare the results obtained from the second and third specifications for the purchase incidence and basket size models, we observe that those from the former specification are more parsimonious than those from the latter. This is because the scale parameter corresponding to the linear effect by the number of participating local vendors in the third specification of the purchase incidence model is not statistically significant (see coefficients for σ_1 and σ_2 in Table 2 for Model 3). Moreover, the value obtained for ψ_1 in the third specification of the basket size model is of very low magnitude (0.0021) and not statistically significant. Finally, the likelihood ratio test reveals that Model 3 does not fit the data better than Model 2 ($P\ value = 0.6193$), making it more appropriate to use the estimates from the second specification in our interpretation of the results.

[9] The unobservables in the utility from purchasing on the platform are positively correlated with the observed price as indicated by the positive estimate of the price control. This is consistent with the literature: because the platform sets the price, the unobservables in this model are typically "appealing" attributes that are positively correlated with consumers' willingness to make a purchase on the platform (Villas-Boas and Winer, 1999). In contrast, conditional on controlling for bias in the price estimate, the basket-size effect is capturing the relationship between price and quantity. Therefore, as shown by the negative estimate of the price control, the unobservables in the basket size are expected to be negatively correlated with price by first principles, namely the law of demand.

Interpreting the Results in the Purchase Incidence and Basket Size Models

In the second specification of the purchase incidence model, the estimated effect of price is negative and statistically significant (-0.3564, $p < 0.01$), reflecting a downward slope of the demand curve with respect to price, as expected. In particular, a 1% increase in price at the platform decreases the likelihood of purchasing by 0.9897%. [10] In addition, the estimated first order effect for the number of local vendors is positive and significant (0.1104, $p < 0.01$). However, the negative and statistically significant estimate for the quadratic term for number of local vendors (-0.0010, $p < 0.01$) suggests that this likelihood is concave in the number of local vendors. Evaluated at the mean values, a 1% increase in the number of local vendors in the platform increases consumers' likelihood of purchase by 1.0232% (based on a marginal effect estimation of 0.5976%).

The estimates of consumers' need-based effects are largely consistent with our expectations. As expected, the likelihood of purchasing at the platform during a vacation week is negative and statistically significant (-0.2006, $p < 0.01$). Furthermore, those consumers previously enrolled in the platform's delivery membership program have a higher likelihood of purchasing at the platform in the future (0.3963, $p < 0.01$). According to the marginal effects obtained, the likelihood of purchasing at the platform decreases by 3.55% in vacation weeks while the choice by consumers to sign up for a delivery membership leads to an increase of 7.83% in this probability.

Moreover, a consumer's historic rate of consumption has a positive and significant effect on the likelihood of purchasing at the platform (0.0740, $p < 0.01$). This implies that, as expected, as a consumer's average consumption rate increases by one unit,

[10] The elasticity of price is calculated as $\frac{\partial Pr(inc)}{\partial p} \times \frac{p}{Pr(inc)}$, where the marginal effect of price is obtained as $\frac{\partial Pr(inc)}{\partial p} = -\alpha Pr(inc)(1 - Pr(inc))$. The elasticity and the marginal effect of other explanatory variables in the purchase incidence model are also calculated in this way.

the probability of this consumer purchasing at the platform goes up by 1.3381%. In addition, the number of items purchased by consumers in their most recent order has a negative and statistically significant effect ($-0.0027,\ p < 0.01$), suggesting that, after controlling for consumers' average expenditure rates at the platform, those individuals who purchased one additional unit in their most recent transaction at the platform are 0.0495% less likely to purchase in week t. However, this negative effect is small relative to the positive marginal effect by consumption rate. Additionally, the estimated purchase interval effect is negative and statistically significant ($-0.0685,\ p < 0.01$), suggesting that for each additional week in the time interval since a consumer's most recent purchase at the platform, the consumer is 1.2387% less likely to make a future purchase at the platform. This result implies that consumers' purchase behaviors at the platform are driven by loyalty such that the longer the time between purchases, the less likely it is that consumers will purchase at the platform in the future.

Turning to the results in the second specification of the basket size model, we find that the estimated effect of price is negative and statistically significant ($-0.0876,\ p < 0.01$), reflecting a downward slope in the demand curve with respect to price, as expected. According to the elasticity value (-0.3190) obtained from the price effect estimate, a 1% increase in price at the platform yields a 0.3190% decrease in basket size. [11] The estimated effect by the number of local vendors participating in the platform ($0.0087,\ p < 0.01$) suggests an elasticity of 0.3548. This implies that a 1% increase in the number of local vendors increases the basket size by 0.3548%.

In addition, we obtained positive and significant estimations for the effects on basket size by (1) the dollar amounts in coupons used by consumers per order, per week and (2) consumers' prior enrollment in the platform's delivery membership program

[11] The elasticity of price is calculated as $\frac{\partial \lambda_{ht}}{\partial p} \times \frac{p}{\lambda_{ht}}$, where the marginal effect of price is obtained as $\frac{\partial \lambda_{ht}}{\partial p} = \lambda_{ht}\psi_p$. The elasticity and the marginal effect of other explanatory variables in the basket size model are also calculated in this way.

(0.0172, $p < 0.01$ and 0.2682, $p < 0.01$ respectively). According to our marginal effect estimations, a one dollar increase in the value of coupons used by consumers yields an increase of 0.4095 items per week in basket size, and consumers' choice to subscribe for delivery memberships increases basket size by 6.2779 units. The magnitude of the latter effect suggests that subscribed consumers use the platform to fulfill a far greater amount of their grocery demand and therefore purchase a larger basket size conditional on the purchase incidence.

Finally, we obtained a positive and statistically significant estimate for the effect on basket size by the number of items purchased by consumers in their most recent order (0.0090, $p < 0.01$). This implies that given the choice of shopping at the platform, a consumer's basket size will increase by 0.2143 units in relation to each additional unit purchased in the most recent order. This effect is different to the one estimated in the purchase incidence model, where the effect is negative and statistically significant (-0.0027, $p < 0.01$), suggesting that those individuals who made larger purchases as part of their most recent transactions at the platform are less likely to purchase in the future. Combined, these opposite effects further imply that a cumulative effect by a consumer's patronage influences her decision on the size of her orders at the platform. This decision differs from that of buying at a traditional grocery store, which depends negatively on the lagged quantity related to consumers' inventory levels at home.

5 .3 Structural Model: Supply Provision Stage

In this stage of the structural model, we quantify the marginal effect on the platform's surplus from expanding the number of participating local vendors while controlling for the simultaneous effects of platform prices and demand on the size of the platform's local vendor base. The demand effect follows from aggregating at the market level the results obtained in the first stage of the structural model. From the analysis in the first stage, we know that expanding the number of local vendors in the platform is conducive to increasing demand. However, for a virtuous cycle of indirect network effects to emerge, local vendors must also value increases in demand harnessed through the platform. Results in the second stage showing a positive marginal effect on the platform's surplus caused by an expansion in the number of participating local vendors would provide evidence, consistent with **Hypothesis 1**, suggesting that local vendors *do* obtain value from increases in demand channeled through the platform.

Because the analysis in the supply-provision stage tests for the effects of price and the number of participating local vendors in equilibrium, we analyzed Equations (11) and (12) jointly while using a GMM estimation with instruments to address potential endogeneity. As pricing and number of local vendors are determined simultaneously, we estimate these two equations jointly. Table 3 presents the results from this estimation approach (GMM with instruments) as well as those from a nonlinear seemingly unrelated regression (SUR) estimation that makes no attempt to account for endogeneity.

Although the signs and statistical significance of the estimated parameters are consistent across both estimation approaches, there are quite a few differences in the magnitudes of the parameters in Equation (11). Therefore, we chose to focus our analysis on the results from the GMM estimation model with instruments since these results account for endogeneity of system equations. Note that because the statistics commonly used to evaluate goodness of fit in regression models (e.g., R^2) are not available for GMM (Olivares and Cachon, 2009), we used Hansen's J statistic to test the specification of this model. The intuition behind the J statistic is that if the proposed specification is a true process of relating the endogenous and exogenous variables, then the conditional moments used in the model must match the observed sample asymptotically. Because this statistic is not statistically significant ($\chi^2 =$

3.2535, $P\ value = 0.1966$), we conclude that our model is correctly identified, and the instruments taken together as a group are valid (Hansen and Singleton, 1982; Mukhopadhyay *et al.*, 1997; Hall, 2005).

Table 3: Supply provision

	Model 1: Nonlinear SUR		Model 2: GMM(IV)	
Average Price	Estimate	Std. Err.	Estimate	Std. Err.
Constant	-2.917***	0.796	-2.109***	0.798
Wholesale Cost	0.953***	0.049	1.034***	0.053
Cost of Utility	0.527***	0.142	0.328**	0.162
Hourly Retail Wage	0.205***	0.058	0.138**	0.058
Retail Margin (Conduct Parameter φ)	0.066***	0.013	0.079***	0.016
Number of Local Vendors				
Constant	-31.754***	2.871	-32.306***	3.110
Marginal Value Local Vendors ($\frac{1}{\delta_1}$)	0.086***	0.016	0.086***	0.018
N	128		127	
$R^2 LLF/G$	0.72		n.a.	

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10%.
Model 1: Nonlinear seemingly unrelated regression.
Model 2: GMM estimation with instruments.

All parameter estimates are of the expected sign. We first refer to the estimation of the φ coefficient. The results show that this coefficient (0.079, $p < 0.01$) is positive and statistically different from zero. This suggests that the platform's pricing behavior deviates from perfect competition and that the platform can leverage consumers' preference for local vendors in pricing. In doing so, the platform absorbs some of the available indirect network effects (Weyl, 2010). We then refer to the estimation of the $\frac{1}{\delta_1}$ coefficient in the model to quantify the marginal value of expanding the number of local vendors participating in the platform. According to the value estimated for this coefficient (0.086, $p < 0.01$), bringing in an additional local vendor to the platform has a marginal value of \$0.086 per item in the platform's assortment. This benefit is substantial since it accounts for 20.47% of the platform's average gross margins per item ($0.42 as shown in the summary statistics). Moreover, based on an average of 27 items per basket purchased at the platform, adding a local vendor increases the profit of an average order by $2.32 ($0.086/item $\times$ 27 items). Therefore, the platform could afford to incentivize a local vendor to join with an average $2.32 bonus *for each order*. Furthermore, because we obtain this estimate in equilibrium, this incentive value would be agreeable to the vendor. In all, these results provide support for **Hypothesis 1**.

To conclude this section, we use the values of the parameter estimates from the pricing and number of local vendor equations in Table 3 to simulate the number of local vendors in a new equilibrium that maximizes the platform surplus. According to our analysis, this number corresponds to 40.58 vendors, which is lower than the number of local vendors (65, per Figure 1) available to participate in the platform. This inequality is the result of non-linearities in the platform's indirect network effects imposed by increases in the number of participating vendors in equilibrium. To raise this number of vendors without detriment to indirect network economies, the platform could improve its search and recommendations capabilities in order to increase the utility consumers derive from local vendor variety so that more vendors may benefit from joining in. This would help alleviate any negative *cross-side externalities* that may affect the demand at the platform when the number of participating vendors grows too large. Alternatively, the platform could increase local vendor participation without eroding indirect network economies by subsidizing its distribution costs with support from government plans, including the USDA Local Food Promotion Program. This would help ameliorate any negative *same-side externalities* on the supply side that may be caused by marginally increasing costs in the number of vendors participating in the platform. The next section presents two counterfactual simulations of these strategies to identify the number of vendors in equilibrium under nonlinear indirect network effects and contrast this number against that obtained in equilibrium under linear network effects as a result of both simulations. The gap in the number of vendors observed after contrasting both equilibria in each simulation constitutes the basis to test **Hypotheses 2a** and **2b**.

5 .4 Counterfactual Simulation

The first simulation addresses *negative cross-side externalities* by increasing consumers' preferences for local food vendors in the basket size model in the demand stage. This may involve improvements in the platform's search and recommendations capabilities focusing on increasing the value consumers have for additional vendor variety, particularly for lesser known vendors. To implement the counterfactual analysis, we simulate datasets by varying by 5%, 10%, 15%, and 20% the parameter ψ_n, representing consumer preferences for the number of local vendors in the basket size model in the demand stage. We then simulate the number of vendors in the new equilibria from these scenarios under nonlinear indirect network effects and compare these numbers with those under linear indirect network effects. The gap between the implied equilibria emerging under linear versus nonlinear indirect network effects determines whether the growth of online platforms for local grocery food is subject to nonlinear indirect network economies due to negative cross-side externalities, per **Hypothesis 2a**.

To simulate the equilibria for the scenarios, we use the structural parameters obtained from the consumer demand and supply provision stages to solve for the number of local vendors and prices simultaneously and compute the implied basket size (Draganska and Jain, 2005; Bonnet *et al.*, 2013). Moreover, the simulation focuses on the periods *after* the number of local vendors has reached 40, the number of vendors observed in equilibrium without any intervention, to ensure that the effects of changing preferences are estimated on stable equilibria under nonlinear and linear indirect network effects, respectively.

Table 4 presents the results from the counterfactual simulation under nonlinear and linear indirect network effects (top and bottom panel respectively). The results under both conditions show that a higher degree in consumer preferences for greater local vendor variety makes it possible for the platform to attract larger basket sizes and to increase the number of participating local vendors. However, only under linear conditions are the size of the baskets and the number of local vendors elastic with respect to the degree of consumers' preferences for local vendor variety. For instance, under linear conditions, a 15% rise in the degree of preference for local vendor variety respectively increases the platform's average basket size and the number of participating vendors by 38.03% and 19.41% (to an average of approximately 71 vendors). On the other hand, this same increase in preferences for vendor variety under nonlinear conditions respectively increases the average basket size and the number of local vendors at the platform by only 9.54 % and 2.27% (to an average of approximately 60 vendors). Based on these differences, we can infer that the growth of online platforms for local grocery food is subject to significant nonlinear indirect network economies due to negative cross-side externalities, per **Hypothesis 2a**.

Table 4: Changes in the number of local vendors and basket size after increasing consumers' preference for local vendors

	Increase in preference parameter (%)	Number of Local Vendors			Basket Size		
		Mean	Mean percent change (%)	Std. Dev. of percent change (%)	Mean	Mean percent change (%)	Std. Dev. of percent change (%)
Under Nonlinear Indirect Network Effects	5	59.20	0.75***	0.20	12.00	3.06***	0.22
	10	59.63	1.51***	0.43	12.38	6.23***	0.46
	15	60.07	2.27***	0.70	12.77	9.54***	0.73
	20	60.5	3.04***	1.06	13.18	12.97***	1.02
Under Linear Indirect Network Effects	5	67.91	14.41	13.21	14.68	25.60	22.01
	10	69.03	16.30	11.12	14.95	30.12*	17.74
	15	70.91	19.41*	10.33	15.52	38.03*	22.64
	20	70.82	19.52***	8.99	15.34	39.9	25.63

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10%.
Mean percentage change in the number of local vendors=$100 \times \frac{N^*_{\Delta\psi_n} - N}{N}$.
Mean percentage change in basket size=$100 \times \frac{E(Q)^*_{\Delta\psi_n} - E(Q)}{E(Q)}$.

The second simulation addresses *negative same-side externalities* on the supply side by offsetting the marginal costs at the platform through the use of public subsidies promoting the provision of local grocery foods. Such subsidies would allow the platform to obtain lower rents from participating vendors without compromising its profitability, for example. To implement the counterfactual analysis, we simulate datasets by reducing the marginal cost of the platform in increments of 5%. Moreover, we focus again on the periods after the number of local vendors has reached 40, the number of vendors observed in equilibrium without any intervention, to ensure that the effects from reducing marginal costs are estimated on stable equilibria under nonlinear and linear conditions. Table 5 presents the effects on the number of local vendors and prices in equilibrium obtained from this simulation under nonlinear and linear indirect network effects (top and bottom panel respectively).

Table 5: Changes in the number of local vendors and prices after lowering marginal costs

	Decrease in Marginal Cost (%)	Number of Local Vendors			Price		
		Mean	Mean percent change (%)	Std. Dev. of percent change (%)	Mean	Mean percent change (%)	Std. Dev. of percent change (%)
Under Nonlinear Indirect Network Effects	5	59.17	0.76	0.61	3.61	-4.44***	0.11
	10	59.57	1.51	1.21	3.44	-8.87***	0.22
	15	59.97	2.27	1.79	3.27	-13.29***	0.33
	20	60.38	3.04	2.35	3.10	-17.72***	0.44
	80	65.69	12.74*	7.73	1.11	-70.52***	1.56
	85	66.18	13.62*	8.10	0.95	-74.91***	1.64
	90	66.69	14.52*	8.46	0.78	-79.29***	1.71
Under Linear Indirect Network Effects	5	67.96	14.58	13.32	3.64	-3.39***	1.25
	10	69.15	16.68	10.98	3.47	-7.82***	1.14
	15	70.67	19.33***	8.89	3.31	-12.20***	1.12
	20	72.12	21.92***	7.16	3.14	-16.59***	1.07

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10%.
Mean percentage change in the number of local vendors=$100 \times \frac{N^*_{\Delta mc} - N}{N}$.
Mean percentage change in price=$100 \times \frac{p^*_{\Delta mc} - p}{p}$.

As shown in Table 5, increases in the number of local vendors are more responsive to the simulated effects under linear than under nonlinear conditions. While the number of participating vendors increases by an average of 19.33% (to about 71 vendors) following a 15% cut in marginal costs under linear conditions, achieving this same increase in vendors would require a decrease in excess of 90% in marginal costs under nonlinear conditions, which is not realistic. Furthermore, under linear conditions, increasingly larger cuts in marginal costs decrease equilibrium prices at an absolute rate below that observed for increases in the number of local vendors. For instance, a 15% cut in marginal cost yields a decrease in equilibrium prices of 12.20% which is well below the 19.33% increase in the number of local vendors obtained from this simulation. This does not occur under nonlinear conditions. In this case, reductions in marginal costs do not increase the number of local vendors at a rate large enough to offset the rate of decrease in equilibrium prices. Put differently, the rate of increase in the number of vendors under nonlinear conditions is not large enough to raise consumer demand in order to offset the reductions in equilibrium prices. These growth differences under linear and nonlinear conditions, provide evidence consistent with **Hypothesis 2b**, suggesting that the growth of online platforms for local grocery food is subject to significant nonlinear indirect network economies attributable to negative same-side externalities.

6 Conclusion

Despite a significant rise in consumer interest in local grocery foods, the supply of these products has remained constrained by inefficiencies that have made consumers' access to them difficult. In this paper, we evaluate the economic viability of a two-sided platform model to contribute a solution to this challenge. According to the findings obtained from this model, consumer demand increases in the number of vendors participating in the platform and vice versa, making it possible for its surplus to expand by simultaneously growing both sides of the platform. This suggests that a platform model constitutes a viable alternative for the commercialization of local groceries to consumers and could offer a solution to consumers' limitations in accessing these products.

We study this phenomenon through an empirical analysis of indirect network economies in an online platform providing consumers access to grocery food sourced from local vendors in the US. Using a structural model, we estimate the strength of participation on each side of the platform based on observations involving matches between grocery food offerings from local vendors and consumers' demand for these items. Our results from the demand side show that consumers' demand is increasing and concave in the number of local vendors participating in the platform. Thus, an increase in the number of local vendors has a positive effect on consumer choice but this effect diminishes as the number of local vendors increases. In addition, our results from the supply side show that, conditional on demand, incorporating one additional local vendor provides positive incremental surplus to the platform in equilibrium. Thus, the expansion of the number of local vendors causes a rise both in consumer demand and equilibrium surplus, which is a clear indication that indirect network economies are present in the growth of this platform.

However, we determined that there is a fixed number of local vendors that imposes a cap on the indirect network economies at work at the platform. This suggests that for online food platforms exploiting indirect network effects, the number of participating vendors mitigates the strength of these network economies. As a result, indirect network economies in these platforms are not inherently linear and do not always continue unabated as the platforms scale up. We also perform counterfactual simulations directed at alleviating these non-linearities so that more local vendors can participate in the platform. Such simulations involve improvements in consumer search and vendor recommendations that could reduce frictions in transactions and thus increase the value of consumers' preferences for variety among vendors. They also involve subsidizing the increasing marginal costs the platform incurs as it expands its supply base. Results from counterfactual simulations show that these strategies can increase vendor variety at the platform without detriment to indirect network economies. However, they also show that the non-linearities in indirect network effects that exist at the platform hamper severely these strategies' effectiveness in increasing the number of vendors participating at the platform.

6 .1 Implications for Practice and Policy

Increasing the scope of the vendor base in a local-grocery platform can make it much easier for the platform to lock in the available consumer base on its other side. This is an important consideration since local-grocery platforms are not subject to the direct network effects often observed among users and developers participating in platforms involving technology products (e.g., software, mobile apps, etc.) and services (e.g., cloud computing). As a result, local-grocery platforms do not have the same competitive lock-in seen in these platforms. Additional implications regarding platform seeding strategies can be gleaned if we contrast a local-grocery platform's effectiveness in attracting vendors by exploiting indirect network economies rooted in consumers' preferences for local foods versus exploiting indirect network economies rooted in the marginal costs incurred as the vendor base expands. According to our evidence, the benefits enjoyed by consumers from local vendor variety appear to outweigh the benefits gained by local vendors from greater consumer demand. In the end, consumers are charged for the benefits of accessing a variety of vendors, but vendors are not. This is why, to expand the number of local vendors, the platform needs to induce a higher willingness to join from consumers.

This can be realized by providing consumers with a higher value in accessing a greater variety of local vendors through improvements in consumer search, vendor recommendations, and other initiatives intended to reduce frictions in matching consumers with vendors' offerings. Traditionally, local-grocery platforms' emphasis on these initiatives has been limited (Berti *et al.*, 2017; Bielaczyc *et al.*, 2020). Many of these platforms have chosen to interact with consumers through uniform interfaces based on subscription-based applications that provide little room for customization (e.g., deliverybizpro.com and localfoodmarketplace.com). Our results show that it is critical for these platforms to expand on these interfaces to implement initiatives designed to reduce search frictions encountered by consumers when navigating through product offerings from local vendors. Implementing these initiatives will alleviate the negative externalities surrounding these platforms and will give more local vendors an opportunity to market their inventories outside traditional local channels and grocery chains.

Another avenue to expand the number of vendors in a local-grocery platform involves the use of government subsidies to offset the marginal costs of product distribution to consumers. Despite a growing interest in using this type of policy to expand local food markets, subsidies made available by the government remain limited. Consider, for instance, that the 2020 USDA Local Food Promotion Program awarded only 17% of its total annual $13.5 million fund in the form of subsidies to help expand six online platforms across the country. Limitations in these amounts may be the result of a lack of awareness among policy makers of the important role that subsidies play in supporting the growth of local-grocery platforms. Our results provide evidence that these subsidies can help these platforms overcome expansion hurdles by offsetting increasing costs in fulfilling consumer demand as these platforms grow their supply base. With these subsidies, platforms are expected to generate stronger indirect network effects and therefore a greater surplus for not only participating vendors but also consumers purchasing on the platforms.

With a greater surplus, a local-grocery platform can sustain its growth even after subsidies end. Opportunities to direct these subsidies to generate this type of growth abound in the US, given that most platforms across the country involve small supply bases. Therefore, it is important for policy makers to evaluate these platforms and prioritize those that will generate the greatest social welfare from these subsidies. Of particular interest are platforms that serve consumers in markets with limited access to local groceries, platforms where vendors have a greater opportunity to differentiate their products, or platforms with high distribution costs.

6 .2 Opportunities for Future Research

Our study provides scholars with several opportunities for additional research. First, concerning trackability, our analysis is based on static Nash equilibria that assume that equilibrium prices and the number of local vendors reflect rational expectations among platform stakeholders. That is, the dynamics are subsumed in their correct expectations as to how the game is going to evolve if played out repeatedly, over time (Benjaafar *et al.*, 2020; Halaburda *et al.*, 2017). Future research could explore platforms' dynamic decisions in pricing and the number of local vendors to assess our results in relation to the effects based on a dynamic setting.

Second, we simulate how platforms can alleviate negative cross-side and same-side externalities through improvements in platforms' search and recommendations capabilities and the use of subsidies. Assuming improvements in platforms' search capabilities are implemented and more policy subsidies do become available for platforms, we expect there will be opportunities for future research to evaluate the impact of these strategies using field data. This evaluation may involve a comparison between different applications designed to improve search versus recommendations for consumers or comparisons in the use of subsidies among vendors across different product categories.

Finally, our data describes the operations of an online food platform over the course of only two-and-a-half years. Future research could study the operations for this type of platform over a longer period of time to analyze the evolution of indirect network economies in the long run. This analysis may offer additional insights into different indirect network effects brought about by the number of participating vendors as platforms become more mature. For example, it is possible that the indirect network effects of adding more vendors vary across product categories in the platform, particularly if the opportunities for product differentiation are significantly greater in some categories than in others. If the benefit of having additional vendors is more pronounced for some categories, then the platform should invest more resources in expanding the vendor base in these categories as they can generate a greater benefit.

Chapter 2

STRUCTURAL ESTIMATION OF DRIVER ATTRITION IN A LAST-MILE DELIVERY PLATFORM

Abstract

We consider the question of how to better manage turnover among independent drivers who transport parcels for last-mile delivery platforms. Although driver attrition in these platforms is both costly and difficult to manage, there is little understanding of the processes responsible for this attrition. We collaborate with a last-mile delivery platform to build a structural model that enables us to estimate the effects of key predictors of drivers' decisions to continue or leave the platform. For this estimation, we apply a dynamic discrete-choice framework in a two-step procedure that accounts for unobserved heterogeneity among drivers while circumventing the use of approximation or reduction methods commonly used to solve dynamic choice problems in the operations domain. Drivers are compensated using a combination of base payments that reward drivers' productivity and supplementary payments that subsidize drivers with subpar productivity. We find that base pay has a greater effect on drivers' retention. Furthermore, the marginal effects of both base and supplementary pay diminish with drivers' tenure at the platform, but the latter diminishes much faster than the former. We also show that drivers' ambivalence between quitting and staying at the platform affects the effectiveness of compensation as a lever to manage attrition. This effectiveness is at its highest when drivers are "on the fence" between leaving and remaining at the platform. In addition, we find significant heterogeneity among drivers in their non-pecuniary taste for the jobs at the platform and a significantly greater

probability of retention among drivers with greater taste for these jobs. Last-mile delivery platforms can leverage our results to improve driver retention. Through counterfactual analyses, we offer recommendations on how platforms can cost-effectively improve retention by changing the allocation of funds between base and supplementary pay.

1 Introduction

Online retail sales have grown rapidly over the past two decades to account for almost 20% of all retail revenue in the U.S. (Keyes, 2017). To keep pace with this growth while maintaining flexibility, a number of last-mile distribution service providers (e.g., OnTrac, Quiqup, TForce, USPack) have resorted to using platform models based on independent drivers who own and operate the vehicles that make deliveries on behalf of online retailing firms. Alas, high demand for these drivers coupled with the flexibility they enjoy in leaving one service provider's platform for another has exposed these organizations to high levels of turnover among this workforce (Straight, 2018). Because drivers act as the providers' brand image in the eyes of customers, not only is losing them expensive, it also negatively affects customers' perceptions of these companies. In addition, excessive driver turnover can put these companies at risk of exhausting qualified driver resources (Mims, 2019).

Given that attrition is a significant problem for last-mile delivery providers, what then are managers at these companies to do? One option is to identify the factors responsible for attrition and then develop policies that would ameliorate this problem. Although prior research (e.g., Steel, 2002; Hom *et al.*, 2017; Moon *et al.*, 2018; Emadi and Staats, 2020) has identified a variety of factors that are important predictors of attrition, it has yet to do so in the context of last-mile delivery service providers operating based on platform models. Moreover, although recent studies (e.g., Cachon *et al.*, 2017; Allon *et al.*, 2018; Liu *et al.*, 2019; Benjaafar *et al.*, 2020) have examined staffing challenges in transportation platforms, they have focused exclusively on platforms devoted to the transportation of passengers and have given no consideration to attrition in those platforms. Through our examination of drivers' attrition processes in a platform devoted to the delivery of online retail orders to consumers, we contribute to addressing this gap in the literature.

As a setting to study attrition, we focus on TForce Logistics, a provider of last-mile delivery services operating a platform of drivers who work as independent contractors delivering parcels annually for a variety of online retailers. Our study benefits from access to a unique dataset and an in-depth understanding of workforce challenges at TForce gained from field visits and interviews with managers, recruiters, dispatchers, and drivers. Historically, attrition among drivers has been a constant challenge at TForce as well as elsewhere in the industry. Annual driver turnover rate across the industry has been increasing for the better part of the last decade to approach the 100 percent mark a couple of years ago (Straight, 2018).

In order to examine attrition at TForce, we model drivers' decisions to leave the platform over time. When drivers join the platform, they bring with them their own vehicles and commit to a weekly schedule to work moving forward in one of the metropolitan areas served by the platform. TForce assigns to every new driver a delivery route available in these areas. Drivers will typically serve only this route during their tenure at the platform. Drivers may leave the platform at any time after the start of their work. The only requirement is that they give TForce a week's notice before their departure. Therefore, drivers will make decisions to leave or stay at the platform on a weekly basis.

The utility of drivers from staying or leaving the platform depends upon their compensation, their effort (i.e., hours driven) necessary to earn this compensation, their length of tenure at the platform, and their non-pecuniary taste for the job. Drivers' compensation at TForce is comprised of a "base" and a "supplementary" payment, received on a weekly basis. The former is a piece-rate payment: it is a function of the number of parcels (or pieces) that drivers deliver along the routes assigned to them when they join TForce. As the number of pieces delivered increases, the base compensation drivers receive for serving the routes increases. Therefore, this compensation rewards with higher earning rates ($/hour driven) those drivers who are more productive in delivering all the pieces assigned to the different stops located along their routes.

Although the routes assigned to the drivers generally remain fixed over the drivers' tenure, exogenous variations in consumer demand induce frequent changes in the number of pieces delivered and in the location of the delivery stops along the routes. Therefore, drivers' earning rates obtained from base compensation will exogenously increase in the number of pieces but decrease in the distance between the stops along the routes. Furthermore, earning rates from base compensation will increase in drivers' tenure at TForce. Because routes assigned to drivers generally remain fixed over the drivers' tenure, the amount of experience that drivers accumulate on the job will contribute to decreasing their effort (i.e., hours driven) in making deliveries. As this effort decreases, earning rates from base compensation increase.

TForce's supplementary compensation is meant to subsidize new drivers as they gain experience and become more productive in their routes. [12] This compensation is set up to lift the earning rates new drivers obtain from their base compensation up

[12] In some instances, supplementary payments also compensate drivers when exogenous variations in consumer demand in their routes impose unusual work requirements in the form of attended deliveries, special handling conditions, long hauls, etc.

to a level comparable to the rates earned historically by more experienced drivers. As driver's tenure increases, TForce decreases the drivers' supplementary compensation. Whether the rate of decrease in this compensation is too slow or too fast will depend on the drivers' non-pecuniary affinity for the job. Ideally, TForce could gauge how long to maintain supplementary payments in place and how high these payments should be while taking into consideration drivers' non-pecuniary taste for the job. However, non-pecuniary taste is *a priori* unobservable and, thus, it is not possible to characterize it before drivers join TForce. Thus, TForce's compensation may ultimately be ineffective in maximizing drivers' retention.

Our goal is to develop a model that will allow us to understand the factors that influence driver retention and evaluate compensation strategies that address the challenges TForce faces in retaining drivers during their tenure at its platform. To that end, we analyze drivers' decisions to leave or stay at the platform via a dynamic choice model in which drivers maximize their utility of staying or leaving over their tenure at TForce. We use a structural model to evaluate the drivers' decisions to stay or leave as a function of their weekly compensation (from base and supplementary pay), the effort they must put in to obtain this compensation every week (as a function of hours and miles per stop driven), their length of tenure at the platform, and their non-pecuniary taste for the job. We estimate structural parameters using nearly six years of weekly payroll data. In addition, we identify the properties of the model, derive comparative statics results, and conduct counterfactual analyses to assess the effectiveness of alternative retention policies.

Our paper offers important contributions to methodology, theory, and practice. Methodologically, our paper contributes the first study in the operations management domain to estimate a dynamic choice model using a two-step framework that makes estimation feasible without having to resort to approximation or reduction methods that are commonly used to solve dynamic choice problems in the literature (i.e., artificially discretizing the state space or approximating the conditional value function). Moreover, we develop our framework using both analytical-based comparative statics and simulation-based counterfactual analyses to evaluate policy implications for retention. To address the challenge of having a large state space, we follow the framework originally proposed by Hotz and Miller (1993) that makes estimation feasible by helping ease computational burden in the analysis without the need to compress the state space. For the application of this framework, we use a two-step algorithm by Arcidiacono and Miller (2011) in order to account for unobserved heterogeneity among decision makers (e.g., TForce drivers). In the first step, we estimate the parameters of the state transition probability functions and the conditional choice probability function. Then, in the second step, we use these results to estimate the structural parameters of drivers' utility from leaving and staying at the platform.

Our paper also contributes to theory relating to job turnover. The labor economics literature has examined factors that influence employee retention, including the seminal paper by Jovanovic (1979). This work presents a theory, based on the "job-matching hypothesis," that helps explain the observed phenomenon that a worker's probability of job separation tends to decrease in tenure. Prior to employment, the match between a worker and a job is subject to imperfect information. As time on the job increases, the quality of the match is revealed, those with a poor match leave the job, and thus the probability of quitting decreases with tenure. We find support for the job matching hypothesis at TForce. In particular, drivers' utility increases with tenure, and consequently, the probability of quitting decreases with tenure. Moreover, our paper builds on this theory by exposing factors that influence the sensitivity of a worker's decision to stay or quit to changes in compensation. We find that sensitivity depends on the worker's non-pecuniary taste for the job as well as the product of two factors: (1) the worker's marginal utility of compensation, and (2) the worker's degree of ambivalence toward a decision to either quit or remain at the platform. It is likely not surprising that marginal utility of compensation matters, though it is important to recognize that different forms of compensation are valued differently. More significantly, however, is the role of ambivalence, which is measured as the product of quit and stay probabilities. Sensitivity is increasing in ambivalence. The key observation is that a worker's ambivalence toward a job, which could be measured and influenced over time by a firm, affects the leverage of compensation for retaining workers. This is a general result that holds when a worker's utility from continuing with a firm is linear in incentives.

Through our estimates from the structural model, we also contribute to practice by exposing a number of managerial insights. First, we find that drivers' non-pecuniary value for the job affects their likelihood of staying at the platform. According to our results, 40.40% of drivers have a high value for the job and therefore are more likely to stay relative to the other 59.60% of drivers who have a low value. Compared to drivers in the former group, those in the latter group receive higher supplementary pay (56.31% higher, on average) and yet have average lengths of tenure that are more than three times shorter. Thus, while supplementary pay may incentivize these drivers to stay at the platform, it may also keep them from learning about their true productivity until later in their tenure, at which point they will decide to leave the platform.

Second, we find that base pay is more important for driver retention than supplementary pay. A $100 increase in weekly base compensation raises drivers' rate of retention by 32.94%, whereas the same increase in supplementary compensation yields only a 20.91% increase in retention. Moreover, we find that this gap is largest for drivers who require lower supplementary payment amounts. Thus, base pay is more effective at increasing retention than supplementary pay, particularly among more productive, less subsidized drivers.

Third, as alluded to above, we find that retention becomes less sensitive to compensation over time. Moreover, because this phenomenon is more pronounced for supplementary pay, it implies that supplementary compensation programs have a greater risk of becoming ineffective for retention if supplementary payments remain too high long after drivers have joined the platform.

Building on these insights, we use counterfactual interventions to evaluate different compensation strategies that address the challenges TForce faces in promoting retention. We show how to improve retention without increasing compensation costs by shifting part of the compensation drivers receive through supplementary pay to base pay. Since drivers value base pay more than supplementary pay when making decisions to stay at TForce, a shift in compensation from supplementary pay to base pay improves retention at the platform. We also show how TForce can maximize retention by not only shifting compensation from supplementary pay to base pay but also by increasing the supplementary payment amounts received by drivers whose supplementary compensation lags that of other drivers in the platform.

2 Literature Review

Our work is related to four streams of literature. The first comprises labor economics studies of worker attrition (see Ashenfelter and Card, 2010, for a review). A prominent hypothesis in this literature argues that a worker's probability of continuing at a job increases in tenure due to a higher value in the matching between the worker and the job (as reflected in greater worker productivity and higher earnings). This hypothesis, commonly known as the "job-matching hypothesis," is rooted in the work of Jovanovic (1979) and has been applied to a variety of business sectors, including the long-distance truckload segment of the for-hire trucking industry (Hoffman and Burks, 2020). Jovanovic (1979) argues that the evaluation of a match between a worker and a job prior to the start of employment is subject to imperfect information. This is because jobs are akin to "experience goods" (in the terminology of Nelson (1970)) and, therefore, the only way for a worker to determine the true value of the match with a job is to experience it. As the worker's experience at the job increases, the arrival of new information about the job will allow the worker to form a more accurate valuation of the match with the job. The worker will leave the job if the value of the match is revealed to be low. Moreover, the probability that this will occur is a decreasing function of the worker's tenure at the job. Our study of drivers' decisions to quit or remain at the platform builds on this work because drivers are forward-looking in making these decisions and maximize their utility of staying or leaving over their tenure at the platform. Therefore, drivers will choose to extend their tenure at their job if their matching value with the job is higher than their utility from quitting the job.

The second area of related work includes studies in the fields of management and psychology looking at factors linked to worker attrition (see Hom *et al.*, 2017, for a review). Among these factors, studies have identified correlates of attrition such as workers' compensation and job demands in a variety of employment settings, including trucking (Shaw *et al.*, 1998; De Croon *et al.*, 2004) and warehousing (Gardner *et al.*, 2011). Rarely, however, have studies in these fields considered these factors jointly as a system to explain workers' decisions to leave or remain at an organization nor have they analyzed dynamical worker attrition phenomena. This is despite calls from scholars for research that addresses this deficiency in the literature (Mitchell and James, 2001; Steel, 2002). Through the development and empirical application of a dynamic choice model explaining the attrition process of drivers during their tenure at a platform, we contribute to fill this gap in the literature.

The third literature stream includes studies that examine compensation schemes that incentivize drivers to vary their work effort in passenger transportation (or "ride-sharing") platforms. For instance, Chen (2016) uses surge-pricing data from Uber to examine elasticity in the number of hours that drivers commit to working at the platform relative to the compensation they receive. The results show there is positive elasticity among drivers. Allon *et al.* (2018) also find support for the existence of a positive elasticity in labor supply compensation using data from a different ride-sharing platform. According to their analysis, drivers are more likely to commit more working hours when offered higher financial incentives. More recently, Guda and Subramanian (2019) show how platforms can strategically use surge pricing to incentivize drivers to work more hours in market zones where the number of drivers is insufficient to serve passengers. Similarly, Liu *et al.* (2019) show how a platform can use bonuses to incentivize workers to allocate more labor hours to the platform at the expense of its competitors. To our knowledge, no study has examined drivers' decisions to leave these platforms permanently. We address this gap in the literature, within the context of last-mile deliveries of online retail orders to consumers.

The fourth stream of literature includes work in operations management that has used structural estimation to manage schedules in healthcare operations (Olivares *et al.*, 2008), model customer attrition in call center queues (Akşin *et al.*, 2013), and model demand by airline passengers (Li *et al.*, 2014) and by consumers in the retail industry (Musalem *et al.*, 2010). Other work has examined labor attrition in business process management (BPM) contact centers and in manufacturing plants (Moon *et al.*, 2018; Emadi and Staats, 2020) using a structural estimation of optimal stopping models similar to the one we use in our paper. We not only apply our model to a new setting and apply a two-step estimation approach that is new to the operations management literature, but also evaluate effects on attrition of different forms of payments over workers' tenure. That is, rather than drawing inferences on attrition from compensations based on fixed hourly or annual payments, we are able to estimate effects by payments that reward workers' productivity (as a function of base compensation) and payments that subsidize workers' efforts (as a function of supplementary compensation).

3 Data, Descriptive Statistics, and Preliminary Evidence

We investigate turnover taking place among drivers involved mainly in the delivery of online orders for Office Depot and Staples. We focused on the delivery of orders for these retailers because they accounted for a large percentage (47%) of pieces delivered by TForce at the time we conducted our study. Additionally, these two retailers have similar route characteristics and consistency in the type of products delivered in the routes as well as among the consumers (e.g., commercial businesses, hospitals, educational institutions) who receive their products. Drivers assigned to these two retailers also operate similar vehicles (delivery vans). Moreover, none of the routes from these two retailers had abnormal profiles, e.g., involved no long-haul services. This gives us a pool of drivers and routes that allows for a more reliable assessment of compensation and efforts in our model. In this section, we first present the details of our dataset along with the descriptive statistics. We then conduct a series of analyses to provide preliminary evidence of our structural model.

3 .1 Data and Descriptive Statistics

Our analysis spans the period between January 1, 2014 and August 31, 2019. The analysis focuses on 15,293 observations from 396 drivers receiving at least 80% of their compensation from deliveries made for Office Depot and Staples across nine U.S. metropolitan areas with different traffic and weather characteristics as well as a range of working conditions in the TForce facilities (i.e., the branches) where drivers pick up their orders for distribution. [13] These drivers represent the majority of workers who made deliveries for Office Depot and Staples during the period of analysis. The drivers worked at TForce for at least three weeks, ensuring we are able to generate a minimum number of observations to simulate each driver's state transition. Moreover, the drivers joined TForce on or after the start of our sample period (January 1, 2014) and the last driver to join did so three months prior to the end of the analysis period (May 25, 2019). This allows us to avoid an initial conditions problem typically observed in dynamic models with unobserved heterogeneity when initial periods are not available in the data as well as to reduce potential biases in our analysis caused by right censoring in drivers' tenure length.

All the deliveries made by the drivers in their metro areas were same-day or next-day deliveries. In each of the metro areas, the drivers arrived at the TForce branch to pick up the parcels for the routes assigned to them when they joined TForce. After picking up the parcels, the drivers then made their deliveries along their assigned routes. In some weeks, drivers did not work due to illness or other personal reasons. Apart from this, their work varied from one day to another depending on the orders required for delivery along their routes by Office Depot and Staples' consumers. The routes assigned to the drivers when they joined TForce generally remained fixed over their tenure at the platform and, within these routes, deliveries involved regular stops at institutional (hospitals, schools, government agencies) and commercial (office buildings, retail establishments) locations that generally remained unchanged over drivers' tenure.

[13] The nine metropolitan areas are: Los Angeles, Orlando, New York, Philadelphia, Houston, Chicago, Detroit, Minneapolis—Saint Paul, and Atlanta. TForce had only one branch in each of these areas during our period of analysis.

Therefore, the variations in the routes centered mainly around the number of pieces and the distances between stops in the routes, and these are determined exogenously by demand from Staples and Office Depot's customers. Appendix B (Part 1) provides evidence that routes' assignments are exogenous with respect to drivers' experience and that they do not depend on the base and supplementary compensations received previously by the drivers.

As we mentioned before, drivers were required to give TForce one week's notice before their departure. Upon drivers' leaving the platform, TForce recorded their date of termination. As shown at the bottom of Table 6, 299 drivers had left the platform by the end of the period of analysis. Of these, 254 lasted 50 weeks or less while 290 lasted 100 weeks or less. Figure 2 illustrates the relationship that exists between the length of time drivers work for TForce and the probability of drivers leaving the company. We observe that the probability of drivers leaving TForce decreases significantly during their tenure. Additionally, drivers' length of tenure appears not to depend on whether drivers joined the platform during its annual peak or regular demand seasons. Specifically, as we show in Part 2 of Appendix B, tenure length appears not to be shorter among drivers who joined TForce during its annual peak demand season (from November to January), when the need for drivers may be more urgent. Furthermore, this part of the appendix provides evidence that drivers' decisions to leave the platform were not likely to occur simultaneously with decisions to leave by other drivers. Therefore, it is unlikely that drivers' decisions to leave affected other drivers' decisions to continue at the platform.

Drivers received compensation from TForce on a weekly basis corresponding to a base and a supplementary payment. The base payment corresponds to a set percentage of base revenue TForce received from the retailers for the routes assigned to the drivers every week. This revenue was negotiated in advance between TForce and the retailers

Figure 2: Probability of drivers leaving TForce

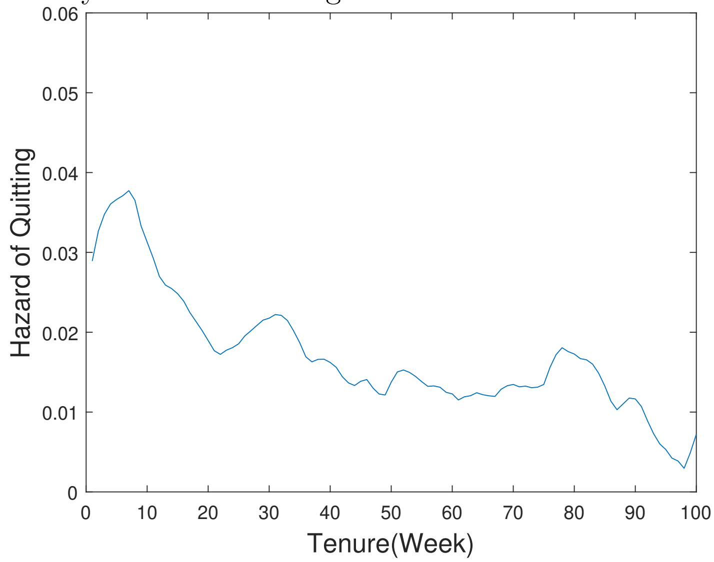


Table 6: Descriptive statistics

Variables (Measured Weekly)	**Mean**	**Std. Dev.**	**Prob. Quit**	**Base Pay**	**Supplement Pay**	**Hours**	**Distance**	**Number Stops**	**Miles per Stop**	**N**
Prob. Quitting	0.02	0.138								15293
Base Pay	821.137	431.313	-0.126***							15293
Supp. Pay	136.938	160.487	-0.037***	-0.068***						15293
Hours	31.521	12.679	-0.136***	0.684***	0.102***					15293
Distance	454.813	345.216	-0.081***	0.426***	0.158***	0.496***				15293
Number Stops	150.942	79.090	-0.132***	0.732***	-0.044***	0.742***	0.453***			15293
Miles per Stop	3.347	2.911	0.017**	-0.003	0.131***	0.003	0.548***	-0.219***		15293
Tenure	45.812	47.867	-0.058***	0.123***	-0.125***	-0.035***	0.025***	0.144***	-0.043***	15293
Tenure (in weeks)	[1,50]	(51,100]	(101,253]							
Number Drivers Quit	254	36	9							
Cummul. Drivers Quit	254	290	299							

whose deliveries made up the routes. It is a function of the number of pieces delivered in the routes. A higher number of pieces yields a larger base compensation. In turn, drivers who are more productive in delivering these pieces will obtain higher earning rates ($/hour driven).

Neither the revenue TForce received from the retailers nor the compensation drivers received from TForce were subject to any sort of dynamic pricing policies similar to those observed in other platforms such as Uber. Nevertheless, TForce did receive from retailers pre-negotiated surcharges for routes that involved deliveries with unusual demands from drivers (in the form of attended deliveries, special handing conditions, etc.). Every week, drivers received as part of their supplementary pay all the surcharges attached to the routes assigned to them.

In addition to compensating drivers who are assigned routes with unusual requirements, supplementary payments included a weekly subsidy that new drivers received as they gained experience. Therefore, these payments decreased in drivers' tenure. TForce had a preset starting value for the weekly subsidy drivers received upon joining the platform. This value corresponded to a set percentage of the weekly base revenue TForce received from the retailers for the routes assigned to the drivers. This percentage decreased during the drivers' tenure at TForce depending on the branch and the month of the year. Upon onboarding, drivers knew that they were subsidized with a supplementary payment set to decrease in tenure. However, they were not privy to the details about the subsidy's starting value or its rate of decrease over time.

Table 6 lists the descriptive statistics for the weekly base and supplementary payments per driver. It also includes the statistics for tenure (in weeks), the probability of quitting every week, as well as the weekly number of hours worked, the distance and miles per stop, and the number of stops serviced by each driver. On average, the probability of a driver quitting in any given week during her tenure at TForce is 2%. During tenure, a driver receives average weekly base and supplementary payments worth \$821.14 and \$136.94, respectively. To earn this income, the driver must work 31.52 hours per week and travel a distance of 454.81 miles to 150.94 locations. The standard deviations in Table 6 also reveal a great deal of variance among weekly payments, number of hours worked, distance traveled, and number of stops serviced by drivers. Although drivers generally made deliveries consistently in the same routes and based on the schedules agreed upon when they entered the platform, their deliveries spanned locations and recipients that changed exogenously over time depending on the recipients' demand. This generated a great deal of variability in the routes assigned to them on a weekly basis, along with the drivers' payments and efforts.

Table 6 also reports the correlation coefficients among the variables in our analysis. As shown by these coefficients, the probability of drivers' quitting has a negative correlation with the length of tenure, consistent with the negative slope observed in Figure 2. Moreover, supplementary pay decreases with tenure. Table 6 also shows that base payment correlates highly with the number of stops in the routes. Recall that base payment is directly a function of the weekly base revenue that TForce receives from the routes assigned to the driver. In turn, this revenue depends on the number of pieces in those routes. Because the number of pieces increases directly with the number of stops in the routes, base payment will correlate positively with the number of stops.

3 .2 Preliminary Evidence of Structural Model and Results

We first use a semi-parametric survival model to investigate the effects of compensation and efforts on drivers' decisions to quit. The advantage of a semi-parametric over a parametric survival model is that it allows us to estimate the effects of explanatory variables without making any assumption on the shape of the baseline hazard (Cox, 1972). The main purpose is to use this model to explore the relationship between base and supplementary pay and the probability of drivers' quitting the platform. We estimate a Cox proportional hazard model of quitting of the form:

$$\lambda(t, X) = \exp(X\beta)\lambda_0(t), \tag{13}$$

where $\lambda_0(t)$ is the baseline hazard. A set of explanatory variables X includes weekly base pay, weekly supplementary pay, weekly hours worked, weekly miles driven per stop, driver's age, metro area fixed effects, and month fixed effects. The variation of the data does not allow us to identify the driver fixed effects, but we cluster the standard errors at the driver level.

Table 7: Estimation results for Cox proportional hazard model

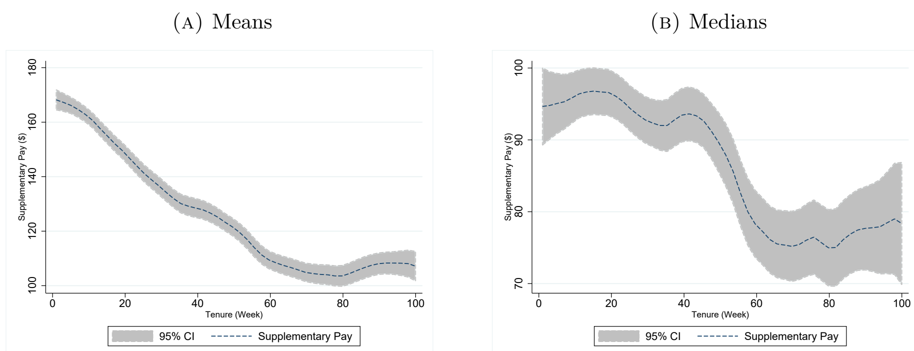
	(1)		(2)	
	Estimate	(Std. Err.)	Estimate	(Std. Err.)
Base pay (week/$100)	-0.343***	(0.033)	-0.349***	(0.034)
Supplementary pay (week/$100)	-0.103**	(0.050)	-0.101**	(0.050)
Hours /10	0.087	(0.086)	0.099	(0.087)
Miles per stop	0.034***	(0.011)	0.033***	(0.011)
Age			-0.007	(0.006)
LL	-2463.918		-2463.131	
obs	15293		15293	
Number of drivers	396		396	

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10.

Columns (1) and (2) in Table 7 present the results obtained after estimating the model with and without driver's age among the explanatory variables. Because these columns yield a similar log-likelihood value and the coefficient for age in Column (2) is statistically non-significant, we interpret our results based on estimates from the more parsimonious specification in Column (1). The estimated effects of both base pay and supplementary pay are negative and statistically significant, suggesting that drivers with lower compensation have a higher chance of quitting. Moreover, the results indicate that the effect of base pay is greater than that of supplementary pay as the estimated absolute value for the base pay coefficient is higher than the absolute value estimated for the supplementary pay coefficient ($|-0.343| > |-0.103|$). In addition, the estimated coefficient for miles per stop is positive and statistically significant, showing that a decrease in route density (i.e., longer distances per stop) increases the chance of quitting.

The semi-parametric survival model only considers the factors that affect the drivers' current utility of quitting. However, drivers' decisions to quit or stay at the platform also depend on drivers' future payouts. Upon onboarding, drivers know that they are subsidized with a supplementary payment that will decrease over time. Figures 3(a) and 3(b) illustrate this phenomenon. To plot these figures, we used the results obtained from a local polynomial regression with an Epanechnikov kernel with a bandwidth of 5 weeks. However, to produce Figure 3(a), we collapse supplementary payments to the weekly mean before using local polynomial smoothing, while to produce Figure 3(b), we collapse supplementary payments to the weekly median. Moreover, in addition to a decrease in weekly supplementary pay, drivers

Figure 3: Decrease in weekly supplementary pay



expect to see an increase in their earning rates ($/hour driven) from the weekly base compensation they receive as their tenure increases and they become more productive. This is illustrated by the results of a polynomial regression of the relationship between tenure and earning rates from base compensation (using weekly means in Figure 4(a) and weekly medians in Figure 4(b)).

Figure 4: Increase in earning rates from weekly base pay

(A) Means (B) Medians

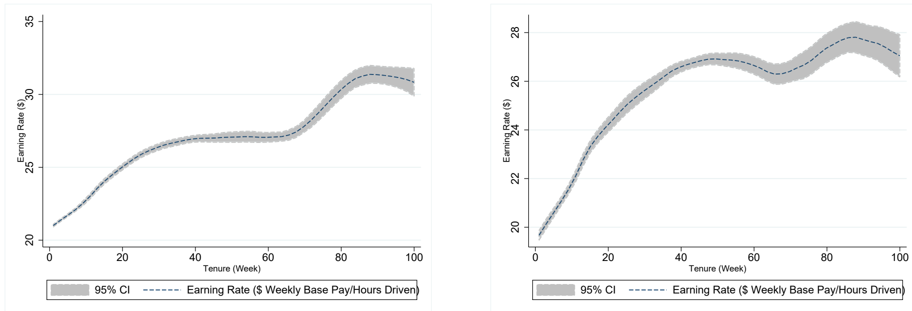


Table 8: Impact of change in compensation on quitting

	Estimate	(Std. Err.)
Percentage change in base pay	-0.006*	(0.003)
Percentage change in supplementary pay	0.001	(0.002)
Base pay (week$100)	-0.230***	(0.044)
Supplementary pay (week$100)	-0.257***	(0.074)
Hours /10	-0.142	(0.114)
Miles per stop	0.038*	(0.023)
Tenure	-0.014***	(0.003)
obs	10,625	
LL	-759.4	
Number of drivers	396	

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10; (4) observations with a percentage of compensation change greater than the 95th percentile are excluded from the analysis.

Building on these driver expectations, we investigate how changes in base and supplementary payments affect drivers' decisions to leave or remain at the platform. To do so, we first calculate the percentage changes in the drivers' weekly base pay and supplementary pay relative to those in the previous week. Then, we use a logistic model to regress drivers' decisions to quit on the percentage changes in their two forms of payment, while controlling for the weekly levels of base and supplementary payments, the weekly amounts of hours and miles per stop driven, metro area fixed effects, and month fixed effects. As shown in Table 8, the coefficient for the percentage of change in weekly base pay is negative and statistically significant, while the coefficient for the percentage of change in weekly supplementary pay is not statistically significant. This suggests that changes in base pay play an important role in drivers' decisions to leave or stay at the platform—drivers who observe an increase in base pay have a lower probability of leaving the platform.

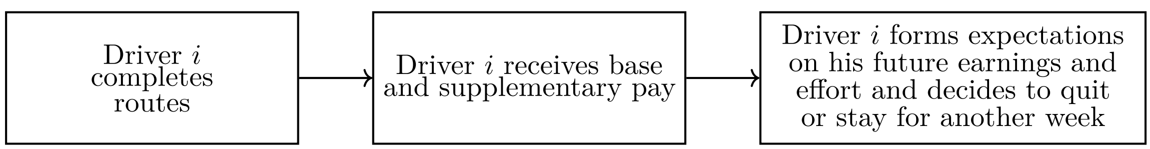
These preliminary findings suggest that drivers make decisions to leave or remain at the platform as a function of their base and supplementary compensation as well as the variations in these amounts over time. Therefore, similar to Chung *et al.* (2014) and Emadi and Staats (2020), we consider drivers as forward-looking decision makers who evaluate their utility of staying at the platform based on not only their current compensation but future compensation expectations.

4 The Model

In this section, we lay out a model for drivers' decisions to leave or remain at the platform. As illustrated in Figure 5, we model these decisions using an optimal stopping framework in a dynamic discrete choice (leave versus stay) context. The standard methods for solving dynamic discrete choice models involve calculating the value function explicitly which can be computationally prohibitive. Conditional choice probability estimators, introduced by Hotz and Miller (1993), provide a tractable alternative to these full solution methods by exploiting the mappings from the value functions to the probabilities of observing particular choices.

As noted in the previous section, drivers are made aware that their payments will change over time and we find empirical evidence that drivers are sensitive to changes in payments. Accordingly, drivers are considered to be forward-looking and take expected future utility into account when making their decisions to quit or stay. By using a dynamic discrete choice model we can account for the possibility that drivers maximize their utility of leaving or staying at the platform over their tenure at the platform, rather than solely during the current period, t. At the end of each week t, an individual driver i working in metro area j must decide whether to stay at the platform for one more week or quit at the end of the week. The decision of a driver is denoted by $d_{it} = k \in \{0, 1\}$, where $d_{it} = 1$ when choosing to stay and $d_{it} = 0$ when choosing to leave. The driver makes the decision by comparing the utility from staying with the utility from quitting over his planning horizon of T periods, as shown in Figure 5.

Figure 5: Decision process at the end of week t



Assuming an additive random utility model, the driver i's per-period utility flow from serving metro area j from decision $d_{it} = k \in \{0, 1\}$ in week t is given by

$$U_{ijkt} = u_k(X_{ijt}) + \varepsilon_{ijkt}. \tag{14}$$

We assume the error term ε_{ijkt} is i.i.d. Gumbel distributed and captures the choice-specific transitory shock. X_{ijt} is a vector of state variables capturing weekly compensation (base and supplementary pay), hours worked, miles driven per stop, and tenure. As reported in Section 5 , our estimation strategy initially assumes that the drivers and the researchers share full information regarding the state variables. It then relaxes this assumption by adding to the model controls for unobserved heterogeneity among drivers.

The individual drivers make the sequence of choices $\{d_{it}\}_{t=0}^{T}$ to maximize the expected present value of utility over the planning horizon of T periods

$$V(X_{ij0}) = \max_{d_{it}} E\left[\sum_{t=0}^{T} \beta^t \mathbf{1}(d_{it} = k)(u_k(X_{ijt}) + \varepsilon_{ijkt})\right], \tag{15}$$

where X_{ijt} is observed in the data while the shock, ε_{ijkt}, is known only to the driver. The driver forms her expectation over the uncertain future states and uncertain future shocks. $\mathbf{1}(d_{it} = k)$ is an indicator function and $\beta \in (0, 1)$ denotes the discount factor. The optimization problem in Equation (15) can be rewritten as a Bellman equation in the recursive form

$$V(X_{ijt}) = \max_{k \in \{0,1\}} \{u_k(X_{ijt}) + \varepsilon_{ijkt} + \beta E\left[V(X_{ijt+1})|X_{ijt}, d_{it} = k\right]\}. \tag{16}$$

Equation (16) completely summarizes drivers' optimal behaviors from week t onward with a per-period utility function and a future expected utility component. The per-period utility from staying at the platform is

$$u_1(X_{ijt}) = \theta_0 + \theta_1 W_{ijt} + \theta_2 I_{ijt} + \theta_3 H_{ijt} + \theta_4 D_{ijt} + \theta_5 T_{ijt} + \theta_{\xi j} + \theta_{\eta t}, \tag{17}$$

where W_{ijt} is the weekly base pay and I_{ijt} is the weekly supplementary pay; H_{ijt} is the number of hours worked in a week; D_{ijt} denotes the route density measured as miles per stop; and T_{ijt} is the week of tenure. Furthermore, we include in Equation (17) the metro area fixed effect denoted by $\theta_{\xi j}$ to control for any effects due to idiosyncrasies in the metro area where the drivers work. A vector of month dummy variables, $\theta_{\eta t}$, captures seasonal effects that may have an impact on the utility of staying at the platform. Because only differences in utility matter and as a standard assumption for the outside option, we normalize the mean utility a driver obtains from quitting as zero and denote $u_0(d_{it} = 0, X_{ijt}) = 0$ (Arcidiacono and Miller, 2011; Arcidiacono and Ellickson, 2011).

For $d_{it} = 1$, we rewrite Equation (16) to denote the value function of staying at the platform in terms of model primitives

$$V_1(X_{ijt}) = u_1(X_{ijt}) + \varepsilon_{ij1t} + \beta E\left[p_0(X_{ijt+1})V_0(X_{ijt+1}) + p_1(X_{ijt+1})V_1(X_{ijt+1}))\right]. \quad (18)$$

Value function, $V_k(.)$, is the driver's state-dependent optimal utility over the remainder of the planning horizon given decision $d_{it} = k \in \{0, 1\}$. The probability that driver i decides to stay at the platform for another week is denoted as $p_1(.)$ and $p_0(.) = 1 - p_1(.)$ is the probability that driver i decides to quit. More specifically, $p_1(X_{ijt+1})$ and $p_0(X_{ijt+1})$ denote the probability that driver i stays or quits at the platform in week $t + 1$ given state X_{ijt+1}. We see in Equation (18) that the driver's optimal value function of staying depends on the future utility evaluated over all possible uncertain states in week $t + 1$, given a decision to stay in week t, and a per-period utility of staying, $u_1(X_{ijt})$, defined in Equation (17).

For $d_{it} = 0$, we rewrite Equation (16) to denote the value function of quitting the platform in terms of model primitives

$$V_0(X_{ijt}) = u_0(X_{ijt}) + \varepsilon_{ij0t} + \sum_{s=t+1}^{T} \beta^{s-t} E(u_0(X_{ijs}) + \tilde{\varepsilon}_{ij0s}). \quad (19)$$

Note that choosing $d_{it} = 0$ implies $d_{is} = 0$ for all week $s > t$ because quitting is a terminating action and, once it is chosen, a driver's decision problem is no longer dynamic. In addition, because we normalize the per-period utility of quitting, $u_0(X_{ijt})$, to zero, the value function for quitting only includes the driver's expected idiosyncratic term over the future periods.

A conditional choice probability estimator makes use of the one-to-one mapping between the difference in choice-specific conditional value functions and the probabilities of making a choice (Hotz and Miller, 1993). Therefore, to derive the conditional choice probabilities estimator, we need to formulate the differenced value functions across two choices as a function of the conditional choice probability. We start by defining the conditional value function as $v_k(X_{ijt}) \equiv V_k(X_{ijt}) - \varepsilon_{ijkt}$. The choice-specific error term ε_{ijkt} is identically and independently drawn from a Gumbel distribution, with the location and scale parameters normalized to 0 and 1, respectively. The difference between two Gumbel random variables is a logistic random variable (Train, 2009), thus

$$p_0(X_{ijt}) = Pr(\tilde{\varepsilon}_{ij1t} - \tilde{\varepsilon}_{ij0t} \leq v_0(X_{ijt}) - v_1(X_{ijt})) = \frac{1}{1 + e^{v_1(X_{ijt}) - v_0(X_{ijt})}} \tag{20}$$

from which we invert to express the difference in the conditional value of each decision as the log ratio of probability of quitting and probability of staying

$$v_0(X_{ijt}) - v_1(X_{ijt}) = \log\left(\frac{p_0(X_{ijt})}{1 - p_0(X_{ijt})}\right). \tag{21}$$

In Appendix C, we further show that the difference in the conditional value functions can be expressed as

$$v_1(X_{ijt}) - v_0(X_{ijt}) = u_1\left(X_{ijt}\right) - \beta \int \log\left[p_0\left(X_{ijt+1}\right)\right] f\left(X_{ijt+1}|X_{ijt}\right) dX_{ijt+1}. \tag{22}$$

Note that the difference in conditional value functions contains only one future component, the one-period ahead conditional choice probabilities, $p_0\left(X_{ijt+1}\right)$, which correspond to the probabilities of choosing the terminal choice, i.e., quitting the platform in week $t+1$. This features a one-period ahead property where the only future value needed is the drivers' probability of quitting evaluated at all possible one-period ahead states (Arcidiacono and Ellickson, 2011). We use Equation (22) in our estimation strategy as described in the next section.

5 Estimation Strategy

In this section, we expand on the strategy we use to estimate the parameters in the dynamic choice model of driver turnover. We use a two-stage approach that obviates the need to solve the dynamic problem repeatedly. In the first stage, we estimate the state transition functions and the conditional choice probabilities. Then, in the second stage, we take these results as given to estimate the structural parameters. [14] Note that although estimating the model in stages does not affect the consistency of the estimates, it does reduce efficiency. We compute the standard errors using subsample bootstrap methods, where we resample the data 200 times to reestimate the model and construct a bootstrap distribution of the estimates (Chung *et al.*, 2014; Murphy, 2018).

5 .1 *Step 1: Conditional Choice Probabilities and State Transitions Estimation*

We use a flexible logit model to estimate the first-stage conditional choice probabilities of the decision to quit. The flexible logit model has been extensively used in the literature to estimate conditional choice probabilities (e.g., Arcidiacono and Miller, 2011; Yoganarasimhan, 2013; Huang and Smith, 2014; Fang and Wang, 2015). The model contains second order polynomials of base pay, supplementary pay, hours worked, density, tenure, and interactions among base pay, supplementary pay, hours worked, density, and tenure as well as metro area dummies and month dummies. This flexible structure is similar to the specification used by Ellickson *et al.* (2012) and is intended to capture as much information as possible to reflect the conditional choice probabilities. Appendix D presents these results.

Next, we estimate the parameters of transition probability functions, $f\left(X_{ijt+1}|X_{ijt}\right)$. Some of the components of the state variables are time-invariant (metro areas) or will transition deterministically (tenure), which simplifies the transition probability

[14] This estimation approach has been extensively used in the marketing and economics literatures (e.g., Misra and Nair, 2011; Chung *et al.*, 2014; Murphy, 2018).

function. We estimate the transitions of the remaining components of state variables following the equations detailed in Appendix D.

5 .2 Step 2: Estimation of Structural Parameters

Finally, we simulate $\int \log\left[p_0(X_{ijt+1})\right] f\left(X_{ijt+1}|X_{ijt}\right) dX_{ijt+1}$ using the coefficients of transition probability functions, $f\left(X_{ijt+1}|X_{ijt}\right)$, the empirical distributions of the state variables, and the conditional choice probabilities, $p_0\left(X_{ijt+1}\right)$. The estimation procedures are detailed in Appendix E. We take $\int \log\left[p_0(X_{ijt+1})\right] f\left(X_{ijt+1}|X_{ijt}\right) dX_{ijt+1}$ as given and estimate the structural parameters of the dynamic choice model, θ, based on the following log likelihood function

$$LL(\theta) = \sum_{i=1}^{I} \log\left[\prod_{t=1}^{T} \mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta)\right]. \tag{23}$$

5 .3 Unobserved Heterogeneity

The likelihood function using Equation (23) is based on the assumption that there are no unobservable preferences among individual drivers. As we discussed before, this is unreasonable because drivers are likely to differ unobservedly in their non-pecuniary preferences for the job. Thus, we assume that there is unobserved heterogeneity in drivers' non-pecuniary taste for the job drawn from a mass-point distribution (Keane and Wolpin, 1997; Heckman and Sedlacek, 1985). This will allow us to explain those cases in which drivers with higher compensations decide to leave early, while drivers with lower compensations decide to stay, for example.

We address the presence of unobserved heterogeneity in drivers' taste for the job by assuming that there are two types of drivers with π_r being the proportion of the rth type in the sample. Types are time-invariant which are known by the individual drivers but are not observed by the researchers. In the estimation, the utility of staying is allowed to differ by an unobserved *Type* variable reflecting persistent non-pecuniary taste for the job. Type enters Equation (17) and the per-period utility of staying becomes $u(X_{ijt}, s_i)$. When a driver belongs to the first category ($r = 1$), type equals 0 ($s_i = r - 1 = 0$), and when she belongs to the second category ($r = 2$), type equals 1 ($s_i = r - 1 = 1$).

The probability of a driver being a particular type r is given by

$$q_{ir} = \frac{\pi_r \prod_{t=1}^{T} \mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta, s_i = r-1)}{\sum_{r=1}^{2} \pi_r \prod_{t=1}^{T} \mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta, s_i)}. \tag{24}$$

The log likelihood function we use to estimate the structural parameters follows the mixture distribution

$$LL(\theta, \pi) = \sum_{i}^{I} \log \left[\sum_{r=1}^{2} \pi_r \prod_{t=1}^{T} \mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta, s_i = r-1) \right]. \tag{25}$$

Note that the parts of the log likelihood function are no longer additively separable. However, Arcidiacono and Jones (2003) show that the Expectation-Maximization (EM) algorithm can reintroduce the additive separability at the maximization step. We use the EM algorithm following the two steps: first, at the expectation step, we calculate the expected log likelihood function given the conditional probabilities at the current parameter estimates; second, at the maximization step, we maximize the expected likelihood function holding the conditional probabilities fixed. We repeat the process until the results converge. As the EM algorithm allows us to return the additive separability, the expected log likelihood function here is given by

$$LL(\theta, \pi) = \sum_{i=1}^{I} \sum_{r=1}^{2} \sum_{t=1}^{T} q_{ir} \log \left[\mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta, s_i = r-1) \right]. \tag{26}$$

We can now proceed with the estimation in stages. In the first stage, we estimate the parameters of the state transition probability functions. Because the transition probability functions do not depend on the type (s_i), they can be consistently estimated in the first stage. In the second stage, we recursively compute and update the conditional choice probabilities, $p_0(s_i = r-1|X_{ijt})$, the population probability of being a type (π_r), and the structural parameters θ. The estimates of the state transition probability functions from the first stage are taken as given to calculate the next period's future continuation values in connection with the updated conditional choice probabilities in each iteration. For the detailed iterations of EM algorithm, we refer readers to Arcidiacono and Miller (2011).

6 Results

Table 9 (Column (1)) reports the structural parameter estimates we obtained after evaluating drivers' decisions while taking into account their unobserved heterogeneity in their non-pecuniary taste for the job among drivers. Appendix F presents an analysis of the model's two-stage estimation performance by comparing the predicted value of quitting with the realized value of quitting observed in the data. This appendix also presents the results from different tests we conducted to evaluate the robustness of the results in Table 9 relative to a variety of discount factors and with respect to a variety of models excluding parameters with non-significant effects in Table 9 and in the conditional choice probability results in Appendix D.

As expected, the estimated parameters for base and supplementary pay effects in Table 9 are positive and statistically different from 0 ($p < 0.01$), while the estimated parameter for the miles per stop effect is negative and statistically different from 0 ($p < 0.01$). Moreover, according to Table 9, the estimated value for the type effect coefficient is positive and statistically different from 0 ($p < 0.01$), indicating that unobserved heterogeneity in the taste for the job affects drivers' decisions to stay. Specifically, compared to type 1 drivers, type 2 drivers have a higher taste for the job and thus are more likely to stay.

Comparing the log-likelihood value obtained for the model specification in Table 9's Column (1) with that obtained for a specification ignoring heterogeneity in drivers' taste for the job (Column (2)), we can infer that the former specification is a better fit for the data than the latter. A comparison of the values estimated for the parameters for all the effects in Column (1) against those estimated in Column (2) also shows a fair amount of consistency, with the exception of tenure's effect. While the values estimated for the parameters for tenure's effect are positive across both specifications in Columns (1) and (2), only the value for the parameter when heterogeneity is ignored is statistically significant. Therefore, tenure's positive effect on retention is contingent on drivers' non-pecuniary taste for the job. This is intuitive, since drivers with longer tenure tend to be the ones who like the job better and are likely to stay longer after controlling for compensation (per the positive and significant value estimated for the parameter for type's effect in Column (1) in Table 9).

Table 10 reports the summary statistics for both types of drivers. The type 1 subsample, consisting of 59.60% of the drivers has, on average, 19 weeks of tenure which is more than 3 times shorter than the 75 weeks of tenure among the type 2 subsample consisting of 40.40% of drivers. Type 1 drivers receive a slightly lower base pay ($808.273 versus $827.173 per week) but work more hours (32.766 versus 30.922 hours). After dividing the base pay by the number of hours worked, we find type 1 drivers receive a lower hourly pay ($24.668 versus $26.750) indicating that type 1 drivers are less productive than type 2 drivers. However, the average type 1 driver receives a higher weekly supplementary pay ($168.300 versus $121.816) while traveling similar miles per stop than the average type 2 driver (3.469 versus 3.289 miles per stop).

These results imply that drivers' non-pecuniary taste for the job plays an important role in their decisions to stay at the platform. The results also imply that TForce's

Table 9: Estimation results with unobserved heterogeneity

	With Type (1)		Without Type (2)	
	Estimate	(Std. Err.)	Estimate	(Std. Err.)
Constant	0.312	(0.353)	0.461***	(0.191)
Base pay (week/$100)	0.127***	(0.030)	0.121***	(0.025)
Supplementary pay (week/$100)	0.094***	(0.035)	0.100**	(0.028)
Hours /10	-0.011	(0.067)	-0.042	(0.057)
Miles per stop	-0.074***	(0.011)	-0.076***	(0.010)
Tenure (week/10)	0.023	(0.017)	0.038***	(0.009)
Type	0.331***	(0.162)		
Type 1 (percent)	0.596			
Type 2 (percent)	0.404			
LL	-1210.024		-1280.307	
obs	15293		15293	
Number of drivers	396		396	

Notes: (1). Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10; (4) The model uses a discount factor of 0.9957 which corresponds to a 0.8 annual discount rate and is consistent with that used previously in the literature (e.g., Hoffman and Burks, 2020).

Table 10: Summary statistics for the two types

	Type 1	Type 2
Type	0.596	0.404
Base pay (week)	808.273	827.173
Supplementary pay(week)	168.300	121.816
Hours	32.766	30.922
Miles per stop	3.469	3.289
Length of Tenure (week)	19.208	75.314

Notes: The population probability of each type is the estimated value for π_r; drivers are classified into a type based on the value estimated for q_{ir} (i.e., driver i is type 1 if $q_{i1} > 0.5$).

supplementary pay may not be as effective in retaining drivers compared to its base pay. As shown in Column (1) in Table 9, the estimated value for the parameter corresponding to the supplementary pay effect is lower than the value estimated for the parameter corresponding to the base pay effect (0.094 vs. 0.127). [15] In Appendix G, we re-estimate the structural parameters in Column (1) assuming the base payments are endogenous and find that the estimated effect of the base payments on driver retention represents a conservative lower bound. In the next two sections, we assess these implications more formally using comparative statics and simulations of counterfactual interventions of the empirical model estimated in the context of our industry partner's operations.

7 Compensation Implications for Retention during Tenure

Our evaluation first considers the extreme where the drivers' discount factor $\beta = 0$. In this case, intuition may suggest that any changes in the state vectors that increase

[15] The estimated values for these two parameters are statistically different at a 0.5% level, according to a likelihood ratio test.

drivers' deterministic utility from continuing relative to quitting and should increase their probability of staying, p_1. The following proposition confirms this intuition (please refer to Appendix H for the proofs for this proposition as well as for those presented later in the paper).

Proposition 7 .1 *Suppose drivers are myopic (i.e., $\beta = 0$).*
For any $x \in \{W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt}, T_{ijt}\}$:

(i) $\dfrac{\partial p_1 (X_{ijt})}{\partial x} = p_0 (X_{ijt})\, p_1 (X_{ijt}) \dfrac{\partial u_1}{\partial x}$

(ii) $\dfrac{\partial p_1(X_{ijt})}{\partial x} > 0$ *if and only if* $\dfrac{\partial u_1}{\partial x} > 0.$

For example, $\dfrac{\partial p_1 (X_{ijt})}{\partial W_{ijt}} = \theta_1 p_0 (X_{ijt})\, p_1 (X_{ijt})$, $\dfrac{\partial p_1 (X_{ijt})}{\partial I_{ijt}} = \theta_2 p_0 (X_{ijt})\, p_1 (X_{ijt})$.

Recall the following notations: $p_1 (X_{ijt})$ is the probability of staying with the platform and $u_1 (X_{ijt})$ is the deterministic component of driver utility from staying. The weekly base pay is denoted as W_{ijt} and θ_1 is its coefficient, while I_{ijt} is the weekly supplementary pay and θ_2 is its coefficient. When $\beta = 0$, all future components in p_1 disappear and only the per-period utility, u_1, remains. Proposition 1 allows for a complete and exact analysis of the effect of changes in any states on driver turnover and the average length of employment.

While Proposition 1 applies to a setting where drivers are myopic, our modeling approach accounts for drivers who are forward looking when making their decisions to leave or stay at the platform. In this case, the discount factor $\beta \in (0, 1)$ and, therefore, the future components are no longer zero. The probability of deciding to continue in the current week, given state X_{ijt} is

$$p_1 (X_{ijt}) = \frac{1}{\left(1 + e^{v_0(X_{ijt}) - v_1(X_{ijt})}\right)}, \tag{27}$$

where $v_k(X_{ijt}) \equiv V_k(X_{ijt}) - \epsilon_{ijkt}$ is the conditional value function. In light of Proposition 1, we may expect that changes in the state in week t leading to an increase in the utility u_1 will cause the probability of staying, p_1, to increase. However, this may not be necessarily the case because a change in X_{ijt} can affect the future states, which in turn can affect the value of $v_1(X_{ijt})$ consisting of not only the per-period utility u_1 in week t but also the expected optimal utility from week $t+1$ onward. Recall that the state variables in week $t+1$, given state X_{ijt} in week t can be expressed as [16]

$$X_{ijt+1} = \left(\varphi_W\left(W_{ijt}, T_{ijt+1}, \tilde{\varsigma}^W_{t+1}\right), \varphi_I\left(I_{ijt}, T_{ijt+1}, \tilde{\varsigma}^I_{t+1}\right), \varphi_H\left(H_{ijt}, T_{ijt+1}, \tilde{\varsigma}^H_{t+1}\right), \varphi_D\left(D_{ijt}, T_{ijt+1}, \tilde{\varsigma}^D_{t+1}\right)\right).$$

If the state transition function φ is non-decreasing in the state variable $x \in \{W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt}, T_{ijt}\}$, then the directional effect of changes in x on the probability of continuing is unambiguous. This result is formalized below. We define $v_1^{(1)} = v_1 - u_1$ (i.e., the value function from period $t+1$ onward conditional a decision to continue in period t).

Proposition 7 .2 *Suppose drivers are forward-looking* (*i.e.,* $\beta \in (0,1)$). *For any* $x \in \{W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt}\}$*:*

(i) $$\frac{\partial p_1(X_{ijt})}{\partial x} = p_0(X_{ijt})\, p_1(X_{ijt}) \left(\frac{\partial u_1}{\partial x} + \frac{\partial v_1^{(1)}}{\partial x}\right)$$

(ii) Suppose $\frac{\partial \varphi}{\partial x} \geq 0$. Then $\frac{\partial p_1}{\partial x} > 0$ if and only if $\frac{\partial u_1}{\partial x} > 0$.

Proposition 2 also indicates that the effects caused by changes in any of the states on the probability of staying depend on the product of p_1 and p_0 and the marginal effects of the state variable on both per-period utility (u_1) and the expected optimal utility from the $t+1$ week onward ($v_1^{(1)}$). Thus, all else being equal, an increase in either

[16] Note that the transition of tenure is deterministic, i.e., $T_{ijt+1} = T_{ijt} + 1$.

base pay or supplementary pay has the greatest effect when drivers are "on the fence" between staying or quitting ($p_1 \times p_0$ is maximized when p_1 is 50%) and the effect is smaller when drivers are more inclined to either staying or quitting. For example, a \$100 increase in base (or supplementary) compensation raises the likelihood of drivers' staying at the platform by 17.02% (or 8.46%) when drivers' probability of staying with TForce is 50%. This same \$100 increase in compensation raises the likelihood of drivers' staying by only 2.48% for base pay and 1.23% for supplementary pay when drivers' predilection to remain at TForce increases to 96.20%. [17] Connecting this result to the two different driver types (in Table 9), we infer that type 1 drivers may be more indifferent between leaving versus staying at TForce than type 2 drivers and, therefore, are more sensitive to changes in compensation. Furthermore, regardless of their types, drivers' indifference between staying or leaving will decrease with tenure as they learn about the matching quality with the job at the platform. Therefore, compensation effects are most pronounced earlier in tenure. A special condition could arise for some unpopular jobs that have a retention rate of less than 50% at the beginning of tenure. In this case, $p_1 \times p_0$ increases first and then decreases in tenure. For these jobs, compensation effects will be more pronounced during the mid-stages of tenure.

From Proposition 2, we can also evaluate the effect of drivers' tenure at TForce as well as the effect of tenure on the sensitivity of retention to compensation. We first evaluate $\dfrac{\partial p_1\left(X_{ijt}\right)}{\partial T_{ijt}}$ at the mean values of all state variables using the conditional choice probabilities in our estimation procedure. Then, to gauge the effects of tenure

[17] These results are obtained from evaluating $\left(\dfrac{\partial u_1}{\partial x} + \dfrac{\partial v_1^{(1)}}{\partial x}\right)$ at the mean values of state variables with the parameters in Table 24. For tenure, we use a value of 3 weeks, corresponding to the shortest duration at TForce among drivers. The probability of staying, p_1, takes either a value of 50% when drivers are "on the fence" or a value of 96.20%, corresponding to the estimated probability of drivers' remaining at TForce in week 3.

on the sensitivity of retention to compensation, we calculate the changes (in percentage terms) in the marginal probability of staying with respect to base pay $\dfrac{\partial p_1(X_{ijt})}{\partial W}$ and supplementary pay $\dfrac{\partial p_1(X_{ijt})}{\partial I}$, respectively, due to an increase in tenure. We find that the marginal effect of tenure is positive. For example, a ten-week increase in tenure leads to an increase of 0.05% in the probability of staying. We also find that the same ten-week increase in tenure lowers the effect of base and supplementary pay on the probability of staying by 7.51% and 19.62%, respectively. [18] Therefore, retention becomes less sensitive to compensation over time.

Having examined the effects of changes in state variables on retention in the myopic and the forward-looking case, we next consider whether the probability of staying, p_1, is more sensitive to changes in one state variable versus another, e.g., base pay vs. supplementary pay. For the case of $\beta = 0$, this type of question can be easily answered using Proposition 1. For example, if $\dfrac{\partial u_1}{\partial x_1} > \dfrac{\partial u_1}{\partial x_2} \geq 0$ for some $x_1 \in \{W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt}\}$ and $x_2 \in \{W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt}\}$, then an increase in x_1 leads to a greater increase in the probability that a driver continues than an increase in x_2. The situation is much more complex for the case of $\beta > 0$. The key reason is that a change in the state variable in the current period affects random future states through the state transition and, in turn, optimal future decisions and the conditional valuation function, v_1, at each realized state. However, under certain conditions it is possible to draw conclusions on the relative sensitivity of p_1 to changes in state variables.

The next proposition applies to a general model with streamlined notation that we summarize below. We suppress the driver (i) and region (j) subscripts, and we let the time period parameter t also denote the tenure of the decision maker (T_{ijt}

[18] To calculate these values, we compare the percentage decrease in the marginal probability of staying from week 45 (the average value of tenure) to week 55 for base and supplementary payments, respectively. Except for tenure, all other state variables are evaluated at their means.

in our model above). The state vector in period t is $X_t = (X_{1t}, \ldots, X_{nt}, t)$ which is comprised of $n+1$ state variables, the first n of which evolve randomly over time via an AR(1) process. The state transition functions are $\varphi_1, \ldots, \varphi_{n+1}$. Given state variable X_{kt} in period t for $k \in \{1, \ldots, n\}$, the random state in the next period is

$$\tilde{X}_{kt+1} = \varphi_k \left(X_{kt}, t, \tilde{\zeta}_{kt} \right),$$

where $\left(\tilde{\zeta}_{1t}, \ldots, \tilde{\zeta}_{nt} \right)$ are iid random variables. The state transition function for time (φ_{n+1}) increments the period by one, i.e., $\varphi_{n+1}(t) = t + 1$. The normalized utility (net of ε) from continuing in period t is

$$u_1 (X_t) = \theta_0 + \theta' X_t \quad \text{where} \quad \theta = (\theta_1, \ldots, \theta_{n+1}),$$

and $u_0 (X_t) = 0$. With this notation, the decision-maker's problem in period t is

$$V (X_t) = \max_{k \in \{0,1\}} \{ u_k (X_t) + \varepsilon_{kt} + v_k^{(1)} (X_t) \},$$

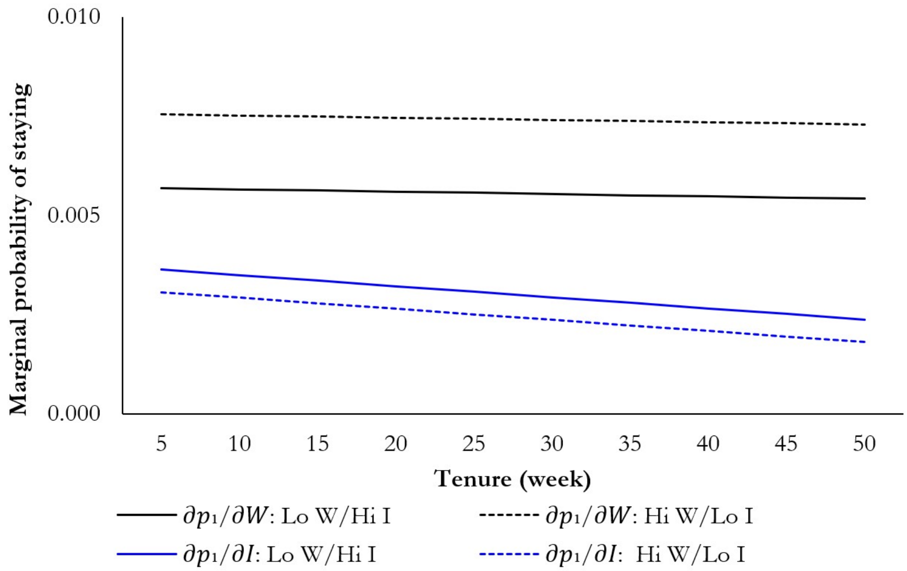
where

$$v_0^{(1)} (X_t) = \gamma \beta \left(\frac{1 - \beta^{T-t-1}}{1 - \beta} \right)$$
$$v_1^{(1)} (X_t) = \beta E \left[v_1 \left(\tilde{X}_{t+1} \right) | X_t, d_t = 1 \right].$$

Proposition 7 .3 *Suppose drivers are forward-looking (i.e., $\beta \in (0,1)$), and the transition functions φ_i for state variables i and j are linear:*
If $\dfrac{\partial u_1}{\partial X_{it}} > \dfrac{\partial u_1}{\partial X_{jt}} > 0$ *and* $\dfrac{\partial \varphi_i}{\partial X_{it}} > \dfrac{\partial \varphi_j}{\partial X_{jt}} > 0$, *then* $\dfrac{\partial p_1}{\partial X_{it}} > \dfrac{\partial p_1}{\partial X_{jt}}$.

Proposition 3 characterizes the relative sensitivity for states with linear state transition functions, as is the case for state variables H and D in our model. The result reveals that the probability of staying is more sensitive to the state variable which has a greater effect on per-period utility and also a higher correlation with its lagged value in the transition function relative to the other state variable. However, this result is not assured to hold for a nonlinear state transition function (e.g., as in W and I). Therefore, we numerically evaluate the sensitivity of retention to base pay and supplementary pay. In particular, we calculate the marginal probability of staying with respect to base pay and supplementary pay being evaluated at two different levels including high and low (denoting respectively 75% and 25% fractiles).

Figure 6: Marginal probability of staying by tenure: base pay (W) and supplementary pay (I)



As shown in Figure 6, the probability of staying is more sensitive to base pay than to supplementary pay. This gap is largest for productive, less subsidized drivers ($\partial p_1/\partial W$: High W/Low I vs. $\partial p_1/\partial I$: High W/Low I) and smallest for highly subsidized and unproductive drivers ($\partial p_1/\partial W$: Low W/High I vs. $\partial p_1/\partial I$: Low W/High I). Therefore, while supplementary pay may incentivize drivers to stay longer in the platform, it may also keep unproductive drivers from learning about their true productivity until later in their tenure, at which point they will decide to leave the platform. Figure 6 also shows that the sensitivity of retention to base pay and to supplementary pay decreases in tenure, with sensitivity to supplementary pay decreasing faster than sensitivity to base pay.

Proposition 4 expands on the results from Proposition 3 by showing how changes in the state transition coefficients affect the probability of retention in subsequent weeks. Let σ_{ik} be a parameter in the state i transition function φ_i. Consider the following condition: $\dfrac{\partial \varphi_i}{\partial \sigma_{ik}} \geq 0$ for all feasible values of the state variables X_{it} in any period t. This condition holds in our model because state variables are nonnegative. In other settings, it may be possible that the sign of a state variable could be both positive and negative (i.e., the sign is not restricted), in which case Proposition 4 does not apply.

Proposition 7 .4 *If* $\dfrac{\partial \varphi_i}{\partial \sigma_{ik}} \geq 0$ *for all feasible values of the state variable* X_{it} *in any period* t*, then the sign of* $\dfrac{\partial p_1}{\partial \sigma_{ik}}$ *matches the sign of* $\dfrac{\partial u_1}{\partial X_{it}}$.

To interpret the implications from Proposition 4, we consider a change in which supplementary pay decreases at a faster rate with tenure. The state transition function for supplementary pay is $\varphi_I = (I_{ijt})^{\sigma_{I2}} \, e^{\sigma_{I1}+\sigma_{I3}(T_{ijt+1})+\zeta_j+\phi_{t+1}+\varsigma^I_{t+1}}$ (see Appendix D) and u_1 is increasing in tenure (see Table 9). Therefore, from Proposition 4, if σ_{I3} decreases (i.e., supplementary pay decreases at a faster rate), then the probability of a decision to stay decreases.

The results from Propositions 3 and 4 suggest that a supplementary payment program can maintain an artificially high level of retention among type 1 drivers with low levels of productivity for the duration of this program. Because these drivers are ill-suited for the job, they will likely end up leaving when the supplementary payment program ends. Unfortunately, it is difficult to identify these drivers before they start their work and, as a result, it is easy for supplementary payment programs to end up allocating money in retaining the wrong type of drivers. According to our propositions, this will be particularly evident in supplementary payment programs that last too long or that pay out excessively high supplements, particularly at the expense of base payments. In the next section, we use counterfactual analyses to examine the insights from Propositions 3 and 4 in the context of our industry partner's empirical model. Our goal is to analyze how TForce could address its attrition challenges as a function of variations in supplementary and base compensations received by drivers for their work at the platform.

8 Counterfactual Analyses

We first evaluate compensation effects on retention by simulating a counterfactual intervention in which drivers' base and supplementary payments increase on a week-by-week basis during drivers' tenure. The goal is to compare the effects on retention caused by increases across both forms of compensation and the evolution of these effects during drivers' tenure. We then simulate a series of counterfactual interventions to provide insights into how our partner could improve its current payment program to increase driver retention as a function of supplementary and base payment costs.

8 .1 Compensation Improvement Effects on Retention

We simulate a counterfactual intervention in which drivers' base and supplementary payments increase by \$100 one week at a time during tenure. That is, both forms of compensation increase by \$100 starting in the third week of drivers' tenure, [19] then in the fourth week, and so on and so forth. The changes in the probability of staying in week t are simulated as $\frac{\sum_{i=1}^{I_t} p_1\left(X'_{ijt}\right) - \sum_{i=1}^{I_t} p_1\left(X_{ijt}\right)}{\sum_{i=1}^{I_t} p_1\left(X_{ijt}\right)}$, where I_t is the observed number of drivers' decisions in week t. X'_{ijt} is equal to the vector X_{ijt}, but with the base pay or the supplementary pay used in X'_{ijt} set to its baseline value used in X_{ijt} plus \$100. Our approach is consistent with that by Kang *et al.* (2015) and Ransom

[19]We start in the third week because no driver in our sample left TForce during the first two weeks of tenure.

(2021) in that it considers interventions that are in effect for only one time period (i.e., one week) at a time. By restricting the interventions to occur only one period at a time, we can recover the counterfactual conditional choice probabilities without solving the full model via backward recursion while benefiting from the computational advantages of the conditional choice probabilities approach (Arcidiacono and Ellickson, 2011).

Figure 7: Compensation improvement effects on retention

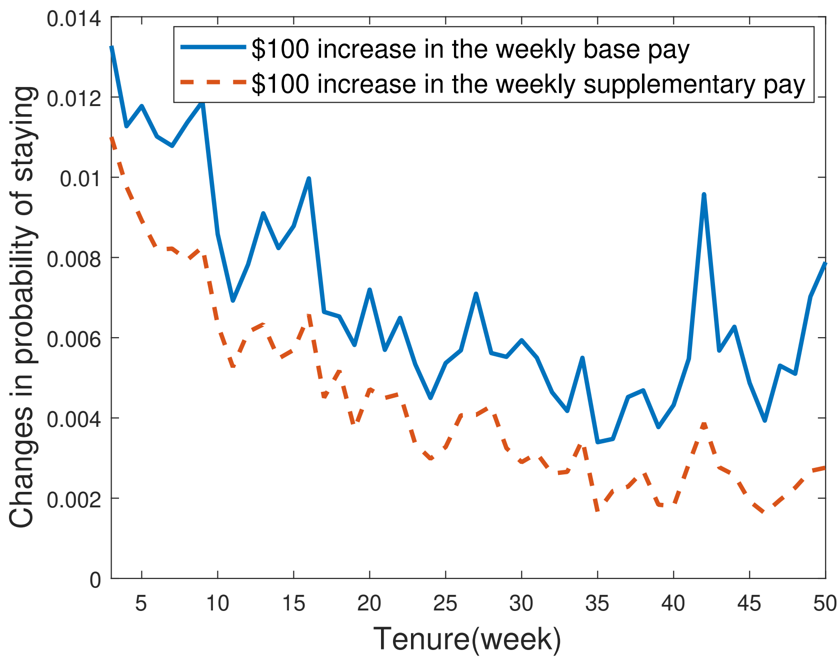


Figure 7 presents the results of the counterfactual analysis for the base and supplementary payment effects on the probability of staying at TForce during tenure. These results show that retention is more sensitive to increases in base pay than to increases in supplementary pay. While drivers' increase in base pay leads to a total increase of 32.94% in the retention rate during tenure (with a weekly average of 0.69% and a standard deviation of 0.25%), the same increase in supplementary pay leads to an overall increase in the retention rate of 20.91% (with a weekly average of 0.44%

and a standard deviation of 0.23%). Thus, in making decisions to stay at TForce, drivers value base pay more than supplementary pay. Moreover, Figure 7 shows that increases in the retention rate caused by higher supplementary and base pay decrease in tenure. Therefore, the impact of base pay and supplementary pay on drivers' utility of staying decreases as tenure increases.

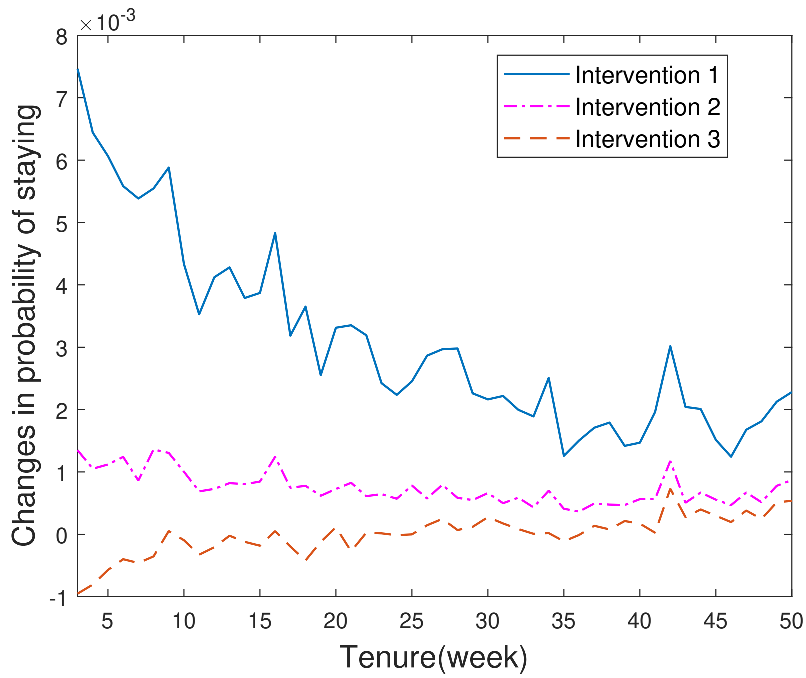
8 .2 Payment Program Policies and Their Effects on Retention

Having examined the effects of supplementary and base payments on retention, we focus on three counterfactual interventions in TForce's compensation program in order to improve driver retention as a function of supplementary and base payment costs. Figure 8 presents the results of the counterfactual analyses of these interventions. In the first intervention, we shift compensation from supplementary pay to base pay and reduce the spread of values of weekly supplementary pay by bringing the two tails of these values' distribution closer to the mean. To that end, we set a target for the weekly supplementary pay at the 60^{th} percentile of the supplementary pay ($120) received by drivers in the third week of tenure. [20] We then calculate the difference between the average supplementary payment drivers receive in their third week of tenure and the supplementary pay target (Δ= $170 - $120 = $50). The value of Δ corresponds to the amount to be added to or subtracted from every driver's supplementary payment received every week, depending on whether this payment amount is below or above the $120 target. Therefore, if a driver's supplementary pay in the fifth week of tenure equals $100, we adjust this amount to $150. However, if this amount equals $180, we adjust it to $130. Moreover, in the latter case, we shift the

[20] We chose this supplementary compensation target for the intervention to ensure that drivers' expectations on their future payments stay close to the primitives in the transition probabilities. We also chose the third week of tenure for this intervention since drivers do not start leaving TForce until after their first two weeks at the platform.

amount subtracted from the driver's supplementary pay to the driver's base pay for that week. In doing so, we increase the emphasis placed by compensation on rewarding drivers' productivity. According to the results from the analysis of this intervention, the cumulative probability of retention at the platform increases by 14.82% during the drivers' tenure. However, such improvement in retention expands compensation costs only by 3.04%.

Figure 8: Results of counterfactual analyses of compensation interventions



The second intervention is a variation on the first one. We again subtract Δ from the driver's weekly supplementary pay when this amount is higher than the target and then shift this surplus to the driver's weekly base pay. However, if the supplementary payment amount is below the target, we leave this amount as well as the base payment amounts unchanged. In the end, this intervention does not change compensation expenses for TForce. However, because it shifts compensation from supplementary to base payments, it increases the cumulative probability of retention during drivers' tenure by 3.57%.

The third intervention is a variation on the second one: It subtracts Δ from the driver's weekly supplementary pay when this amount is higher than the target but does not shift this surplus to the driver's weekly base pay. Instead, the platform retains this surplus. As a result, the compensation costs go down by 7.75%, but the cumulative retention probability decreases greatly by 4.33%, which makes the 60$^{\text{th}}$ percentile an impractical supplementary target for this intervention. Therefore, we instead increase the value of the target to the 75$^{\text{th}}$ percentile in order to obtain a decrease of 3.11% in compensation costs without undermining retention.

Table 11: Summary of compensation strategies and counterfactual analysis results across interventions

	Supplementary Target as a Percentile of 3$^{\text{rd}}$ Week Supplementary Payments	Shift Weekly Surplus in Supplementary Pay to Weekly Base Pay	Eliminate Weekly Deficit in Supplementary Pay	Change in Cumulative Retention Rate	Change in Compensation Costs
1	60$^{\text{th}}$ Percentile ($120)	Yes	Yes	+ 14.82%	+ 3.04%
2	60$^{\text{th}}$ Percentile ($120)	Yes	No	+ 3.57%	0.00%
3	75$^{\text{th}}$ Percentile ($280)	No	No	0.00%	-3.11%

Table 11 summarizes the insights from the counterfactual analyses of the three interventions. Essentially, the second intervention provides the option of improving retention by 3.57% without increasing compensation costs. On the other hand, the third intervention provides the option of reducing compensation costs by 3.11% without eroding drivers' retention rate. The first intervention constitutes an intermediate option as it improves retention by 14.82% while *also* increasing compensation expenses by 3.04%. For a platform facing challenges in retaining drivers, the latter intervention is appealing because it offers a higher improvement in retention than that obtained from the other two interventions. The downside of this intervention is that it carries with it greater compensation costs. A question that arises is whether these additional expenditures are justified by the improvement in retention they afford. On one hand, a 3.04% increase in compensation expenses will increase TForce's payroll by $279,100 over the course of one year. [21] On the other hand, a 14.82% increase in the cumulative probability that drivers will stay at the platform will save TForce \$521,223 annually in recruiting costs for new drivers and in the costs incurred from subsidizing these drivers through supplementary payments as they climb their learning curve. [22] Therefore, because this intervention offers TForce a yearly rate of return of 86.75% (\$521,223/\$279,100 = 1.8675), it is clearly beneficial for the organization.

9 Conclusion

Research on transportation platforms in the operations management literature has focused mainly on how to balance supply and demand in these systems. As such, much of this work has considered topics ranging from worker allocation and scheduling policies, to the management of incentives, to the design of transaction mechanisms. This paper expands this research in a new direction by considering drivers' decisions regarding their continuity in these platforms. In so doing, it contributes to the field a better understanding of the dynamics involved in drivers' behavioral responses to their working conditions in these systems.

[21] To obtain the \$279,100 value, we calculate the difference between the total amount TForce spent on compensation as shown in the data and the amount it would have spent under the intervention during the first 50 weeks of drivers' tenure.

[22] We obtain the \$521,223 value in four steps. First, we determine the number of additional drivers TForce would have retained by the 51st week of tenure due to a 14.82% increase in its retention rate. This number corresponds to 37 drivers and is equal to the product between 14.82% and the number of drivers who had left the platform by the end of week 50 under normal conditions (254 drivers in Table 6). Second, we estimate the costs saved from not having to recruit new drivers to replace the 37 additional drivers retained. These costs are approximately 16% of each driver's annual base payment of \$41,057 (50 weeks $\times$ 821.14/week, per Table 6) or \$6,569.12 per driver (\$243,057 for all 37 drivers). The 16% share is consistent with that used to estimate turnover costs in other studies (Boushey and Glynn, 2012). Third, we estimate the costs saved from not having to subsidize through supplementary payments any new drivers to replace the 37 additional drivers retained. Per driver, these costs are equal to the average supplementary payment observed over the first 50 weeks of tenure (\$7,518/driver). Across all 37 drivers, these costs are equal to \$278,166. Finally, we add the total cost saved in recruiting costs from not having to recruit new drivers to replace the 37 additional new drivers retained (\$243,057) and the costs saved from not having to subsidize through supplementary payments any of these new drivers during their first 50 weeks of tenure (\$278,166) to obtain \$521,223 in savings.

Through our collaboration with TForce Logistics, a provider of last-mile delivery services operating a platform connecting drivers, online retailers, and consumers, we study drivers' decisions to leave or remain at the platform throughout the length of their tenure at the platform. To perform this evaluation, we build a structural model that incorporates into a dynamic discrete choice framework several key predictors of these decisions. These include multiple forms of monetary compensation (through base and supplementary payments), the efforts by drivers to earn these compensations, the length of drivers' tenure at the platform, and the non-pecuniary taste drivers have for the job.

Our model provides an integrated structure to study driver turnover in last-mile delivery platforms. We consider not only the compensation that drivers obtain for their productivity, but also the subsidies they receive as they ramp up their productivity during their early stages of tenure. Moreover, the dynamic choice framework we use in our model enables us to examine the strategic behavior of drivers in their decisions to quit or continue working at the platform. Specifically, our analysis goes beyond the utility drivers obtain in the current period from their choices to leave or stay at the platform to consider their expected future utility of these choices (a forward-looking behavior).

We find that higher compensation, including both base and supplementary pay, increases drivers' probability of staying at the platform. However, compared to supplementary pay, base pay has a greater effect on drivers' retention. While a \$100 increase in weekly base compensation raises drivers' rate of retention by 32.94%, the same increase in supplementary compensation yields only a 20.91% increase in retention. Moreover, we find that this gap is largest for drivers who require lower supplementary payment amounts. Thus, base pay is more effective at increasing retention than supplementary pay, particularly among more productive, less subsidized drivers. These findings are important because they provide a foundation to design compensation programs that are more effective at retaining drivers. Traditionally, research on platforms has not considered the different forms of payments we examine in their design of workers' compensation programs. Our research shows this can lead to erroneous assumptions about the effectiveness of compensation as a lever to increase retention.

We also show that drivers' ambivalence between leaving and staying at the platform affects the effectiveness of compensation as a lever to manage attrition. The effectiveness of base and supplementary pay in contributing to driver retention decreases as drivers develop a stronger predilection toward either leaving or staying at the platform. This decrease in effectiveness is significant. When drivers are indifferent between leaving and staying, a \$100 increase in base (or supplementary) compensation will raise the likelihood of drivers' staying at TForce by 17.02% (or 8.46% in the case of supplementary pay). However, when the bias for quitting or staying increases to 96.20% from the indifference point, this same increase in compensation will raise the likelihood of drivers staying at the platform by only 2.48% (for base pay) and 1.23% (for supplementary pay). This finding is important because research on platforms has typically assumed that the compensation workers receive for their contributions is independent of their degree of ambivalence between leaving and staying at the platforms.

Our results also underscore the role of drivers' tenure in their attrition process. It is known that work experience affects operational outcomes, which in our case involve drivers' decisions to leave or stay at the platform. According to our findings, the probability of drivers' leaving the platform decreases in their tenure. Moreover, the effect of base and supplementary pay on retention decreases in tenure and this rate of decrease is faster for supplementary pay. While a ten-week increase in tenure boosts the rate of retention by 0.05%, it also lowers the effect of base and supplementary pay on retention by 7.51% and 19.62%, respectively. Thus, retention becomes less sensitive to compensation over time. Moreover, because this phenomenon is more pronounced for supplementary pay, it implies that supplementary compensation programs have a greater risk of becoming ineffective at retention if supplementary payments remain too high long after drivers have joined the platform.

Our study also puts a structure around the effects on attrition introduced by drivers' unobserved non-pecuniary taste for the jobs at the platform. We find that about 40.40% of drivers have a high unobserved non-pecuniary value for these jobs and thus are less likely to quit, whereas 59.60% have a low unobserved value. Compared to drivers in the former group, those drivers in the latter group receive higher supplementary pay (56.31% higher, on average) and yet have average lengths of tenure that are more than three times shorter. Thus, the fit that exists between drivers and jobs is an important predictor of attrition at the platform. Our model disentangles the effects of this predictor in the context of TForce.

Finally, we perform counterfactual analyses to arrive at different recommendations on how TForce could address its attrition challenges as a function of variations in supplementary and base compensations received by drivers for their work at the platform. We show how TForce could improve drivers' retention rate by 3.57% without increasing compensation costs or how it could reduce compensation costs by 3.11% without eroding the retention rate. We also show how TForce could increase its retention rate by a significantly higher margin (14.82%) at minimal additional compensation costs (3.04%). This work can serve as an empirical blueprint for other platforms facing similar challenges as TForce's. It can also serve as the basis for studies in other platforms that rely on the crowdsourcing of labor and where the attrition of this labor carries significant downsides including higher recruiting costs and lost productivity.

Chapter 3

AN ANALYSIS OF OPERATING EFFICIENCY AND PUBLIC POLICY IMPLICATIONS IN LAST-MILE TRANSPORTATION FOLLOWING AMAZON'S VERTICAL INTEGRATION

Abstract

We examine how Amazon's decision to vertically integrate its retail platform and last-mile delivery operations can lead to anti-competitive outcomes as a result of a deterioration in the operating efficiency in the routes served by a last-mile transportation firm. Based on an operational analysis of the last-mile transportation firm, we find that Amazon's decision to vertically integrate increases significantly the mileage necessary to deliver parcels in the ZIP code areas where this integration occurs. Moreover, this increase is significantly amplified by the remoteness and proportion of fast deliveries in these areas. These effects translate, on average, into $1.36 in additional costs necessary to cover extra vehicular and labor expenditures per parcel delivered. Because at the root of these outcomes are interactions among multiple organizations with significant market power asymmetries, we expand on a variety of potential anti-competitive service and pricing outcomes stemming from the impact of Amazon's vertical integration on the last-mile delivery firm's costs. We then put forth different public policy remedies that could be implemented to address these sources of anti-competitiveness in the last-mile delivery industry.

1 Introduction

Because of its scale and labor intensity, last-mile transportation represents one of the costliest functions performed in Internet retail supply chains. This has prompted several major retailers to seek greater control over this function by vertically integrating it into their e-commerce platforms. For instance, in 2017, Target acquired last-mile delivery firm, Shipt, for $550 million and rolled out its service across its stores and those of other retailers nationwide (Target Corp., 2017). In that same year, Amazon launched its own in-house service to deliver its own orders as well as the orders of third-party sellers, which account for over 60% of sales in its platform (Kenney and Zysman, 2020).

These vertical integration moves have given rise to a vigorous debate in academia (Bamberger and Lobel, 2017; Borsenberger *et al.*, 2018) and industry (Jansen, 2019; Berman, 2020) over these retailers' misuse of market power in the last-mile transportation industry. While the public policy literature has traditionally considered horizontal integration in the e-commerce industry as anti-competitive, it has provided mixed insights on vertical integration's competitive effects (Khan, 2016, 2019). On one hand, vertical integration may generate benefits, such as lower transaction costs and the elimination of successive monopolies or oligopolies and the double marginalization these entail. However, vertical integration may deprive competing firms of essential inputs, which could decrease their efficiency and preclude them from contesting the market altogether.

This issue is particularly relevant in the last-mile transportation industry. Major retail platforms, particularly Amazon's, have a significant market power in their focal industry (i.e., e-commerce), which gives them monopsony-like power over last-mile delivery firms. Consider, for instance, that e-commerce deliveries account for over 60%

of all parcels transported annually by last-mile transportation firms in the United States and orders from Amazon's retail platform contribute to generating almost half of this volume (Laseter *et al.*, 2018; Duggan, 2020). With such a dominant position, Amazon can deprive last-mile transportation firms of essential delivery volumes and/or cherry-pick its deliveries to decrease the operational efficiency and profitability of these firms (Borsenberger *et al.*, 2018).

In this study, we seek to examine the operating effects of Amazon's 2017 decision to integrate its retail platform and last-mile delivery operations and the implications these effects may have on the last-mile transportation industry. As part of this integration, Amazon shifted the fulfillment of its third-party sellers' orders away from last-mile transportation firms that had traditionally provided this service and put it under the domain of Amazon's own in-house service. Given their dependence on Amazon to market and stock their products, third-party sellers had little say on this move. At the same time, the move effectively deprived last-mile carriers of the stream of parcels from third-party sellers, thereby undermining the operational efficiency in their routes.

Our examination of this erosion in efficiency focuses on the reductions in density experienced in one of these carriers' routes as a result of Amazon's vertical integration. In the transportation economics literature, marginal decreases in outputs (e.g., tonnage, passengers, and parcels) relative to inputs (e.g., mileage) over a given network of routes are reflective of lower density and, therefore, lower efficiency (Braeutigam, 1999). As such, in our setting, we evaluate changes in the number of miles drivers must travel to deliver each parcel in their routes before versus after Amazon's vertical integration. An increase in the mileage per parcel delivered following Amazon's integration will be reflective of lower density in the routes and, consequently, lower efficiency.

We evaluate parcel deliveries across 664 ZIP code areas in the United States. We find that, on average, route densities in each of these areas decreased by 88 percent following Amazon's vertical integration. That is, the average number of miles drivers traveled to deliver each parcel in these areas increased 88 percent once Amazon merged its retail platform and last-mile delivery operations. We also find that this phenomenon is more prominent among routes located in hard-to-reach ZIP code areas. On average, the reductions in density caused by Amazon's vertical integration expand by 0.859 percent for every 1 percent increase in the distance necessary to reach these areas. We also find that reductions in density after Amazon's vertical integration depend on the speed of service consumers' demand across the areas where the routes are located. Reductions in density are 5.45 percent higher for every 1 percent increase in the share of fast (i.e., same day) deliveries relative to slower deliveries in these areas.

These findings have important implications regarding the effects that Amazon's vertical integration may have on competition in the last-mile transportation industry. In the system we analyzed, an 88 percent increase in the mileage required for parcel deliveries translates, on average, into \$40.93 in additional costs necessary to cover extra vehicular and labor costs per route (which corresponds to an additional \$1.36/parcel). In areas that are more expensive to serve (either because of their remoteness or speed of service), these additional costs are even greater. To incentivize Amazon not to vertically integrate and thereby prevent an increase in their costs, last-mile transportation firms, like the one in our study, may opt to decrease the fulfillment rates they charge Amazon (or increase the rates charged to Amazon's retail competitors). This cross-subsidization strategy, however, can distort competition by giving a powerful platform like Amazon's an advantage, since Amazon will pay significantly lower rates than those paid by its retail rivals (Dobson and Inderst, 2008). Another strategy may involve limiting the speed of service offered by the last-mile delivery carrier in the areas where Amazon's integration occurred in order to ameliorate the negative impact of this integration on the carrier's operating costs. This strategy, however, can also distort competition by making it more difficult for Amazon's retail rivals to hire last-mile carriers offering high speeds of service. Our paper expands on these potential anti-competitive outcomes.

Ultimately, these outcomes follow from interactions among multiple organizations with significant market power asymmetries, often found in transportation systems, including the one analyzed in this study. Because existing government policies have been mainly concerned with ensuring that prices in these systems remain low and outputs high, they have not been as involved in overseeing market power dynamics that may ultimately be responsible for these prices and outputs (Khan, 2016). Based on our results, we conclude the paper with suggestions for initial policy improvements that address this oversight within last-mile transportation systems, in general.

2 Literature and Theory Background

As an industry, last-mile transportation has been the subject of wide-ranging research in operations management (OM). This research has examined routing (Gurvich *et al.*, 2019; Voccia *et al.*, 2019; Vareias *et al.*, 2019), dispatching and scheduling operations (Agatz *et al.*, 2011; Yang *et al.*, 2016), revenue management (Agatz *et al.*, 2013; Cao *et al.*, 2018), service quality (Rabinovich and Bailey, 2004; Rabinovich, 2007), and the use of infrastructure for consumer interactions (Han *et al.*, 2019) and order consolidation (Rougès and Montreuil, 2014; Castillo *et al.*, 2018; Deng *et al.*, 2021). Little research, however, has studied last-mile transportation operations in the context of online retail platforms. In this context, research has mainly focused on the crowdsourcing of drivers for deliveries (Qi *et al.*, 2018; Fatehi and Wagner, 2021; Ta *et al.*, 2018). To our knowledge, no studies have evaluated vertical integration effects leading to potential competition concerns involving last-mile transportation firms in those platforms. Our paper contributes to addressing this scarcity in the literature.

Our paper also contributes to the growing literature on online platform design. Prior research in this literature has investigated critical design factors that affect online platform performance, such as market thickness and matching efficiency (Nikzad, 2017; Li and Netessine, 2020; Bimpikis *et al.*, 2020), pricing mechanisms (Bai *et al.*, 2018; Cachon *et al.*, 2019), subsidy policies (Benjaafar *et al.*, 2020; Allon *et al.*, 2018), and governance (Jiang *et al.*, 2011; Zhu and Liu, 2018; Parker and Van Alstyne, 2018). We focus on platforms in the Internet retail industry. A platform in this industry creates value by facilitating interactions among different groups of economic agents (i.e., sellers, consumers, and last-mile delivery operators) located on three different sides of the platform. Greater participation by agents on one side of the platform increases the value of the platform to agents on another side (e.g., Rochet and Tirole, 2006; Armstrong, 2006; Evans and Schmalensee, 2016). As a platform grows in size, however, its pursuit of vertical integration strategies may give rise to potential anti-competitive concerns (Crémer *et al.*, 2019; Shapiro, 2019).

Recent research on vertical integration and firm boundaries in the strategic management and public policy literature has considered this phenomenon (Zhu and Liu, 2018; Wen and Zhu, 2019; Nooren *et al.*, 2018). However, this has not been the case in the OM domain. OM studies have mainly focused on analyzing how vertical integration can ameliorate hold-up concerns imposed by incomplete contracts (e.g., Rabinovich *et al.*, 2007; Park and Ro, 2011; Perols *et al.*, 2013; Steven *et al.*, 2014). In our context involving Amazon's platform, these concerns are generally less salient since most last-mile transportation firms are smaller than Amazon. Moreover, these firms cannot rely on vertical integration to stop a dominant platform like Amazon's from offering similar services. Therefore, if Amazon vertically integrates its platform, it will likely obtain a more favorable position to exploit its power and strengthen its competitive position in the market where last-mile transportation firms compete. (Alimonti *et al.*, 2020).

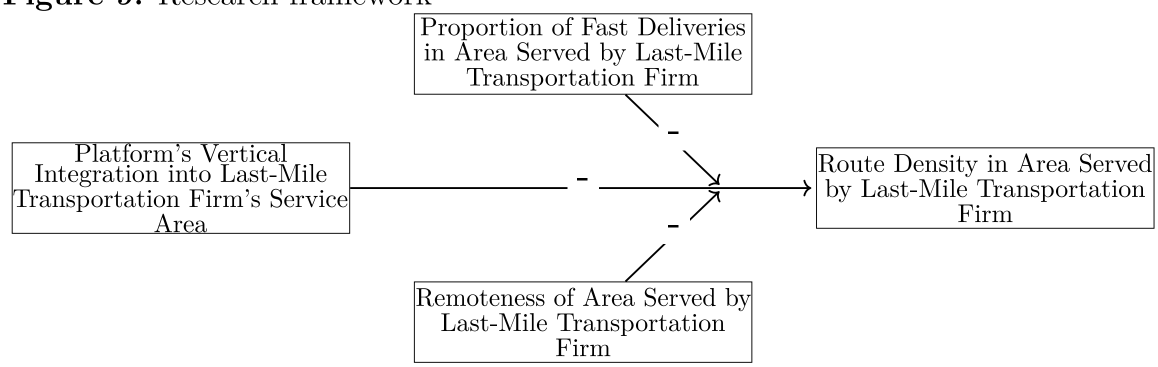
These arguments are based on economic structuralism theories of leverage and foreclosure. While the former theory argues that a firm can use its dominance in one market to extend its power to another (Kessler and Stern, 1959), the latter contends that the firm can use its power in one market to actively deny rivals in an adjacent market access to an essential input originating in the former market (Rey and Tirole, 2007). Therefore, by vertically integrating into its last-mile delivery side, a platform like Amazon's could relegate transportation firms to see route densities decrease in the markets they serve. Eventually, those firms may choose to compete only in those markets where route densities still make it possible for them to operate efficiently – albeit with only those orders that are outside the purview of the platform (Borsenberger *et al.*, 2018). This would leave entire markets subject to less competition in the provision of last-mile delivery services to consumers, thereby compromising the ability of competing retailers or other platforms to contest demand in those markets and ultimately undermining consumer welfare (Alimonti *et al.*, 2020; Zurel and Scorca, 2020; Singh *et al.*, 2021). Moreover, as these markets become less attractive to last-mile delivery firms, drivers may see a reduction in their range of employment and salary opportunities.

We expect that routes affected by vertical integration will exhibit lower densities in markets comprising increasingly remote geographical areas. Relative to other locations, these areas require drivers to travel longer distances (deadheads) to reach them. This leaves shorter amounts of time available to assemble routes with high delivery counts and still meet the time constraints imposed by the delivery dates promised to consumers. A similar situation is evidenced in areas where deliveries must take place under tight time constraints. Consider, for instance, same-day deliveries.

Because this type of delivery carries tighter time constraints relative to other delivery types (e.g., next day), routes in areas with a higher proportion of fast deliveries will exhibit lower densities.

Thus, as formulated in the hypothesis statements below and illustrated in Figure 9, the decision to vertically integrate a platform into its last-mile delivery side may translate into lower densities in the routes operated by a last-mile delivery firm in those market areas where the integration occurred. Furthermore, we expect that this decrease in density will be significantly greater in markets comprised of hard-to-reach areas and areas with a greater proportion of fast deliveries.

Figure 9: Research framework



Hypothesis 3: *A retail platform's vertical integration into its last-mile delivery side decreases the densities in the routes operated by a last-mile transportation firm in the areas where the integration occurred.*

Hypothesis 4a (b): *The remoteness of (proportion of fast deliveries in) the areas served by a last-mile transportation firm amplifies the negative effect on the densities of the routes in those areas caused by a retail platform's vertical integration into its last-mile delivery side.*

3 Empirical Methodology

To perform our empirical evaluation, we must first identify a setting conducive to generating a valid estimation of changes in the density of routes across different areas in the United States before versus after Amazon's move to tie its retail platform and last-mile delivery operations in 2017. To this end, we collaborate with a last-mile delivery firm that operated in various areas across the country that experienced Amazon's vertical integration in that same year. In the remainder of this section, we expand on our research setting and provide details about our data, empirical design, and modeling specification.

3 .1 Empirical Setting

In 2017, the last-mile carrier delivered parcels for more than 200 retailers, including Costco, IKEA, Office Depot, Staples, as well as Amazon. The firm offered two types of deliveries (same-day and next-day) for all retailers. The data we obtained spans all deliveries carried out by the firm across 7,737 ZIP code areas in the United States. [23] These are among the wealthiest and most populated areas in the country (there are 34,097 ZIP code areas in the U.S. where the firm in our study could operate). U.S. household wealth and population follow a Pareto distribution across ZIP code areas, where approximately 20% of the ZIP code areas account for 72% of total household wealth and 68% of total household population. Naturally, last-mile delivery firms like the one in our study, will focus their operations on those ZIP code areas that are wealthier and more populated because this is where they can assemble routes with higher density (parcels/mile). The remaining areas are typically served by firms like the U.S. Postal Service.

Our study focuses on 664 ZIP code areas where Amazon vertically integrated in 2017. These areas account for 70% of Amazon's weekly parcels delivered by the

[23] Please note that the firm does not perform deliveries for all retailers in every ZIP code area it served. In addition, the number and the mix of retailers requiring deliveries in each ZIP code area vary over time.

last-mile firm's drivers in 2017. [24] They also account for approximately 30% of the total weekly number of parcels delivered and 40% of the total weekly number of stops made by the firm's drivers across all 7,737 ZIP code areas served by the firm prior to Amazon's vertical integration. As shown in Table 12, the ZIP code areas where Amazon vertically integrated are distributed across 7 different regions. Table 12 also shows the dates when Amazon vertically integrated in each of these regions. Our focus is on evaluating the effect of these events on the density of the routes across the ZIP code areas in these regions.

Table 12: Regions, ZIP code areas, and Amazon's vertical integration dates

Regions	Amazon Vertical Integration Date	Number of ZIP Code Areas
San Diego	11-Feb-17	62
East Los Angeles	18-Feb-17	152
West Los Angeles	18-Mar-17	101
Chicago	25-Mar-17	143
Central Florida (Orlando and Tampa)	06-May-17	117
Sacramento	19-Aug-17	49
Stockton	18-Nov-17	40

3 .2 Data and Variable Measurements

For each region, we assembled a panel of daily delivery routes collected from the last-mile transportation firm. The firm allocated parcels to these routes based on their delivery addresses and promised delivery time windows. It then assigned each route to a driver working on-demand as an independent contractor and paid by the number of parcels delivered in the route. In total, the panel we assembled includes 1,725 drivers

[24] We excluded from our analysis 124 ZIP code areas where the firm's drivers made deliveries for Amazon in 2017 but did not experience Amazon's vertical integration until 2018 (18 areas) or had yet to experience Amazon's integration as of 2019 (106 areas). We also excluded 177 ZIP code areas where Amazon vertically integrated in 2017 but in which the last-mile delivery firm performed deliveries for Amazon only sporadically prior to vertical integration. We cannot include these areas in our analysis because we are unable to use them to register reliable measurements over time.

and 2,184 routes across the 664 ZIP code areas in our sample. It accounts for 28% of the total drivers (1,725 out of 6,225) and 37% of the routes available (2,184 out of 5,858) in the firm in 2017. For each route, we obtained the name of the driver responsible for the route, the address of the facility where the driver picked up the parcels for delivery (henceforth referred to as the hub), the address and time of delivery for each parcel following the sequence of deliveries, and the delivery service type for each parcel (same-day delivery or next-day delivery).

We used these data to measure the outcome variable (route density) and the moderators (proportion of fast deliveries and remoteness) for each ZIP code area every week. To measure route density, we first estimated the total number of parcels delivered and the miles traveled in every route (from hub to final delivery in the route). We then identified the number of parcels delivered and mileage traveled in each ZIP code area visited by every route. Note that the mileage in each ZIP code area included the deadhead distance traveled to reach the ZIP code area (either from a hub or from another ZIP code area).

Finally, we aggregated across routes the mileage traversed and the number of parcels delivered in each ZIP code area every week. As stated in Equation (28), the ratio of these two measures corresponds to the average weekly mileage per piece [25] delivered along the routes in each ZIP code area. This ratio ($y_{i,t}$) is the reciprocal of the route density for ZIP code area i in week t. [26] As $y_{i,t}$ increases, route density decreases.

$$y_{i,t} = \frac{\sum_{l=1}^{l=7} \sum_{k=1}^{k=K} \sum_{n=1}^{n=N} d_{knl}}{\sum_{l=1}^{l=7} \sum_{k=1}^{k=K} p_{kl}}, \tag{28}$$

where, for a route k in ZIP code area i on day l in week t, d_{knl} is the distance to reach delivery stop n from the previous stop in the route. Since d_{knl} measures the distance of

[25] The terms “piece” and “parcel” are used interchangeably in the rest of the document.

[26] We drop the underscripts t and i from the right hand side of Equation (28) for notation simplicity.

the incoming arc to delivery stop n from the previous stop in route k, it includes both the deadhead distance traveled to reach ZIP code area i and the distance between stops within ZIP code area i. Finally, p_{kl} is the total number of parcels delivered in route k in ZIP code area i on day l in week t.

We measured the weekly proportion of fast deliveries in each ZIP code area ($FD_{i,t}$) by taking the ratio between the number of same-day delivery parcels in the area and the total number of parcels delivered in the area every week. We then measured remoteness for each ZIP code area on a weekly basis ($RM_{i,t}$) by first calculating the distance between the centroid of the ZIP code area and the hub where each of the routes visiting the area originated every week. We then took a weighted average of these distances using as weights the percentage of parcels delivered by each hub to the area during that week:

$$RM_{i,t} = \sum_{h=1}^{H} \left(\frac{p_h}{\sum_{h=1}^{h=H} p_h} \right) d_h, \tag{29}$$

where, d_h is the distance between the centroid of the ZIP code area i and the hub h. [27]

3 .3 Empirical Design

For our analysis, we use a quasi-experimental design with a treatment application corresponding to Amazon's decision to vertically integrate its last-mile deliveries across the different ZIP code areas in our sample. The ideal design to identify the effects of this treatment application would randomly assign Amazon's vertical integration to ZIP code areas in an experiment and track the responses of the outcome variable ($y_{i,t}$) over time. Because such an experiment is obviously unrealistic, we used a quasi-experimental design instead to compare the ex-post responses to the treatment

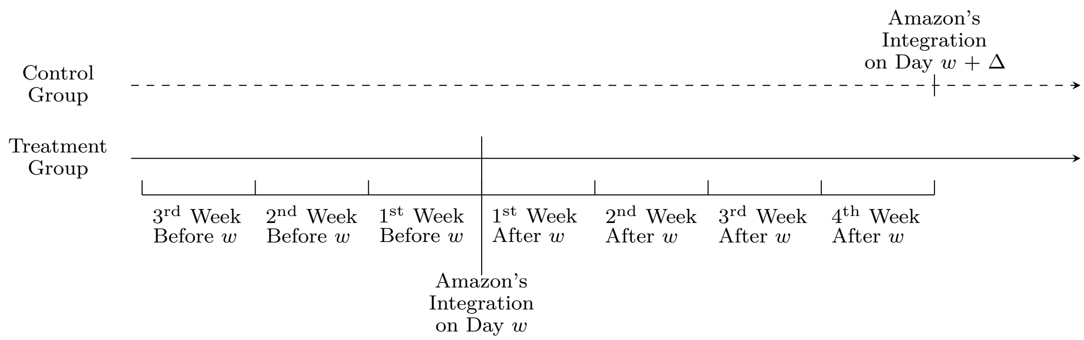
[27]We drop the underscripts i and t from the right hand side of Equation (29) for notation simplicity.

application of affected ZIP code areas to counterfactual ex-ante responses in similar but unaffected ZIP code areas. The goal is to compare ZIP code areas with similar expectations over the distribution of future paths, but with different realizations due to the timing of the treatment application.

To that end, we construct counterfactuals to *treated* ZIP code areas using *control* ZIP code areas that experience the same treatment but a few weeks in the future. As such, our two experimental groups consist of a treatment group composed of ZIP code areas that experience Amazon's vertical integration first, on day w, and a matched control group composed of ZIP code areas that experience the same treatment soon after, on day $w + \Delta$.

This "look ahead" approach has been used by other authors to overcome selection biases in assembling control and treatment groups in quasi-experimental designs similar to ours (Hwang and Park, 2016; Fadlon and Nielsen, 2019; Lim *et al.*, 2021). Since this approach enables us to include in the control group those ZIP code areas that experience integration right after the ZIP code areas in the treatment group, the areas in the former group differ from those in the latter mainly in the timing of Amazon's integration. Figure 10 illustrates this approach. Consider first Amazon's integration in the ZIP code areas in San Diego and East Los Angeles. This integration constitutes the treatment application (on day w). In our analysis, we will evaluate route density in the ZIP code areas in these regions during the three weeks immediately before the date when the treatment application took place (February 11 in San Diego and February 18 in East Los Angeles) and the four weeks immediately after. We will then compare the change in route density in these ZIP code areas before versus after the treatment application against the change observed in route density in the ZIP code areas in the control group during the same period of time. The ZIP code areas in the control group correspond to those in West Los Angeles and Chicago. Amazon's vertical integration in these regions did not occur until after March 18, 2017.

Figure 10: Quasi-experimental design



We follow the same approach for our analysis involving the ZIP code areas in the other regions in Table 12. To study the effect of Amazon's integration in the ZIP code areas in West Los Angeles and Chicago on March 18 and 25, respectively, we evaluate route densities in the ZIP code areas in these regions during the three weeks immediately before March 18 and March 25 and the four weeks immediately after. We then compare the change in route densities in these ZIP code areas before versus after these two dates against the change observed in route density during the same period of time in a control group made of the ZIP code areas in Central Florida, Sacramento, and Stockton where Amazon's integration occurred after May 6. We then perform the same analysis for Amazon's integration on May 6 in the ZIP code areas in Central Florida region by using a control group made of the ZIP code areas in Sacramento and Stockton where Amazon's integration occurred after August 19.

3 .4 Modeling Specification

To evaluate the effects caused by Amazon's decision to vertically integrate its last-mile deliveries, we specify a difference-in-differences (DID) estimator that captures the impact of Amazon's integration on route density based on the following regression specification:

$$y_{i,t} = \alpha_i + \tau_t + \sum_{r \in \mathcal{R}} \gamma_r integration^r_{i,t} \times treat_i + S_{i,t} + \epsilon_{i,t}. \tag{30}$$

In this regression, $\epsilon_{i,t}$ is an error term reflecting Gaussian random shocks on the dependent variable, $y_{i,t}$. Recall that this variable corresponds to route density for ZIP code area i in week t. α_i is a vector of ZIP code area fixed effects. τ_t is a vector of week fixed effects. $S_{i,t}$ denotes the weekly number of stops in each ZIP code area. We include this covariate to account for time-varying amounts in the number of deliveries and drivers across ZIP code areas. $treat_i$ denotes an indicator for whether a ZIP code area belongs to the treatment group. We define $integration^r_{i,t}$ as a dummy variable relative to Amazon's integration date w:

$$integration^r_{i,t} = \begin{cases} 1, & \text{if } t \text{ is } r \text{ weeks after integration} \\ 0, & \text{otherwise} \end{cases}, \tag{31}$$

where, $r < 0$ denotes the relative week leading to Amazon's integration and $r > 0$ denotes the relative week lagging after Amazon's integration ($r \in \mathcal{R} = \{-3, -2, 1, 2, 3, 4\}$). We use $r = -1$ as the baseline. Consistent with other applications (e.g., Lim *et al.*, 2021; Burtch *et al.*, 2018) using the two-way fixed effects approach, we cannot estimate the first-order effects for $treat_i$ and $integration^r_{i,t}$, because after including the ZIP code area and week fixed effects, the coefficient for $treat_i$ is absorbed by the ZIP code fixed effects and the coefficient for $integration^r_{i,t}$ is absorbed by the week fixed effects. The key parameters of interest are γ_r which estimate each week's treatment effect relative to the week leading up to Amazon's integration ($r = -1$). A benefit of using this specification is that it enables us to examine the parallel trends assumption (Angrist and Pischke, 2008). If the treatment and control groups are comparable in the absence of the shock, these groups will follow a parallel trend before Amazon's integration takes place in the treatment ZIP code areas. In this case, we expect to see non-significant effects for all γ_r when $r < 0$.

To quantify mean treatment effects, we use the following standard DID equation:

$$y_{i,t} = \alpha_i + \tau_t + \delta integration_{i,t} \times treat_i + S_{i,t} + \epsilon_{i,t}. \tag{32}$$

In this regression, $integration_{i,t}$ is a binary indicator for Amazon's integration and it takes 1 in all the weeks following Amazon's integration and 0 otherwise. The parameter δ represents the average treatment effect of Amazon's integration on ZIP code areas' outcomes.

Finally, we augment the DID model in Equation (32) by including the moderating factors in $Z_{i,t}$ and interacting them with the treatment effect:

$$y_{i,t} = \alpha_i + \tau_t + (\delta_0 + \delta_1 Z_{i,t})\, integration_{i,t} \times treat_i + \delta_z Z_{i,t} + S_{i,t} + \epsilon_{i,t}, \tag{33}$$

where $Z_{i,t}$ is the average weighted distance between each starting hub to the centroid of the ZIP code area i in week t when studying the remoteness moderator ($RM_{i,t}$ in Equation (29)) or the proportion of fast deliveries in ZIP code area i in week t when studying the fast deliveries moderator ($FD_{i,t}$). The moderating effect on route density is captured by the main parameter of interest δ_1, which is referred to as the difference-in-differences-in-differences (DDD) estimate (Gruber, 1994; Lim *et al.*, 2021; Babar and Burtch, 2020). Note that the first-order effect for the moderator, $Z_{i,t}$, is identified because it varies by week and ZIP code areas and, therefore, it is not absorbed by ZIP code area or week fixed effects. This effect is captured by δ_z.

Finally, we expand Equation (33) by including the two moderators simultaneously and integrating them with the treatment effect:

$$y_{i,t} = \alpha_i + \tau_t + \left(\delta_0 + \boldsymbol{\delta_1}^\top \boldsymbol{Z}_{i,t}\right) integration_{i,t} \times treat_i + \boldsymbol{\delta_z}^\top \boldsymbol{Z}_{i,t} + S_{i,t} + \epsilon_{i,t}, \tag{34}$$

where $\boldsymbol{Z}_{i,t}$ is a (column) vector of moderators which include remoteness ($RM_{i,t}$) and proportion of fast deliveries ($FD_{i,t}$). The moderating effects are captured by a (column) vector of coefficients $\boldsymbol{\delta_1}$.

4 Estimation Results

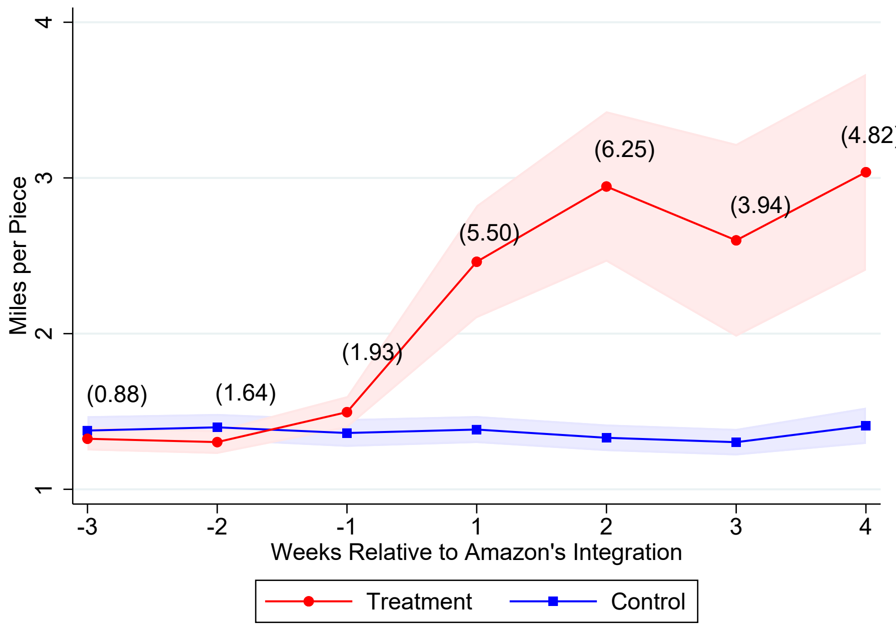
Table 13 presents the descriptive statistics for our key variables during the weeks before and after the treatment application. Before the treatment application, the control and treatment ZIP code areas exhibit similar route density averages (1.44 vs. 1.32 miles per piece). After the treatment application, the average route density in the control ZIP code areas experienced no significant change. The average mileage per piece registered in the routes in these ZIP code areas was 1.41 miles per piece (compared to 1.44 miles per piece prior to treatment). On the other hand, the average route density registered across the treatment ZIP code areas decreased significantly. The average mileage per piece observed across these areas went from 1.32 miles per piece prior to treatment to 2.40 miles per piece after treatment. Moreover, the results from a two-sample t-test during the weeks after treatment show that the average mileage per piece registered for the routes in the treatment areas is significantly higher than the average mileage per piece registered for the routes in the control areas (t-statistic=8.39). Therefore, the route density in the treatment areas became significantly lower than the route density in the control areas after Amazon's integration.

To visualize the reduction in route density in response to Amazon's integration, we plot the average measures recorded for route density for the control and treatment ZIP code areas during each week before and after the treatment application. As shown in Figure 11, mileage per piece immediately increases for the treatment ZIP code areas and the differences between treatment and control ZIP code areas due to this increase are consistently significant across the 4 weeks after the treatment application.

Table 13: Summary statistics for treatment and control groups

	Treatment Group				Control Group			
	Before Integration		After Integration		Before Integration		After Integration	
Time-Variant Variables	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev
Average Mileage per Piece	1.32	1.04	2.40	5.34	1.44	1.12	1.41	1.16
Number of Stops	231.48	212.12	80.22	116.89	277.12	212.04	244.83	205.29
Remoteness (Miles)	16.26	8.41	16.05	8.87	16.95	8.29	16.59	8.38
Proportion of Fast Deliveries	0.33	0.21	0.23	0.27	0.39	0.19	0.37	0.20

Figure 11: Effects of Amazon's integration on route density



The absolute values of the t-statistics comparing route densities between treatment and control ZIP code areas for each week are reported in parentheses.

We expand on these preliminary results by testing for the existence of parallel trends prior to the treatment application. The goal is to verify that route densities in the control and treatment ZIP code areas evolved in parallel prior to Amazon's decision to vertically integrate in the latter areas. We perform the parallel trends test using Equation (30) and report the results in Table 14. The table displays the estimates for γ_r, the vector of coefficients for the interaction between the treatment application indicator and the number of weeks relative to the treatment application. Recall that we measured the time relative to the treatment application using indicators for the number of weeks before and after Amazon's integration, from -3 to $+4$ weeks. The estimates for γ_r capture week r's treatment effect relative to the week right before integration ($r = -1$). Column (1) includes the estimates obtained while controlling for ZIP code area fixed effects and week fixed effects (α_i and τ_t). The estimates in Column (2) are obtained while additionally controlling for the weekly number of stops in each ZIP code area ($S_{i,t}$). Since, as indicated in Table 14, the R^2 value is higher for the specification in Column (2), we will focus on this specification for the interpretation of the results.

Table 14: Relative time model of the effects of Amazon's integration on route density

	(1)		(2)	
	Estimate	Std. Err.	Estimate	Std. Err.
3^{rd} week before Amazon integration (γ_{-3})	-0.039	(0.067)	-0.046	(0.071)
2^{nd} week before Amazon integration (γ_{-2})	0.047	(0.060)	0.033	(0.058)
1^{st} week after Amazon integration (γ_1)	1.020***	(0.180)	1.071***	(0.225)
2^{nd} week after Amazon integration (γ_2)	1.221***	(0.225)	1.278***	(0.265)
3^{rd} week after Amazon integration (γ_3)	1.013***	(0.260)	1.070***	(0.313)
4^{th} week after Amazon integration (γ_4)	1.264***	(0.260)	1.313***	(0.301)
Observations	7,798		7,798	
R-squared	0.3361		0.3362	
Week FE	Yes		Yes	
Service area FE	Yes		Yes	
Control ($S_{i,t}$)	No		Yes	

Notes: 1. Robust standard errors are clustered at ZIP code level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specification in Column (2) includes as a predictor the weekly number of stops in each ZIP code area (S_{it}).

Prior to Amazon's integration, the estimates for γ_r in Column (2) are statistically non-significant, indicating that estimates of route density in the control and treatment ZIP code areas are not different systematically from each other. However, after integration, these estimates become positive and statistically significant. This suggests that the average mileage per piece across the treatment ZIP code areas increased after Amazon's integration relative to the average across the control ZIP code areas. Put differently, route densities in ZIP code areas affected by Amazon's vertical integration decreased in relation to the densities in those areas unaffected by this integration.

Based on the value estimated for γ_1, drivers traveled, on average, 1.071 more miles for every piece delivered in those routes located in treatment ZIP code areas during the first week after Amazon's vertical integration. The values for γ_2, γ_3, and γ_4 suggest that the increases in miles per piece delivered were 1.278 miles in the second week, 1.070 in the third week, and 1.313 in the fourth week. Using Table 13's overall average mileage per piece for the treatment ZIP code areas before Amazon's integration as the baseline (1.320 miles/piece), these mileage increases translate into 81.14% (1.071/1.320), 96.82% (1.278/1.320), 81.06% (1.070/1.320), and 99.47% (1.313/1.320) weekly reductions in route density, respectively.

Table 15 reports the DID estimates for the treatment application effect on route density using Equation (32). The table reports these estimates following the same approach used in Table 14. The results obtained are consistent with those in Table 14. The value estimated for δ across the two specifications is positive and statistically different from zero, suggesting that the average number of miles per piece delivered in the routes located in the treatment ZIP code areas increased significantly after Amazon's vertical integration in those areas. On average, these routes experienced an increase of 1.173 miles per piece delivered after Amazon's vertical integration. This corresponds to an 88.86% (1.173/1.320) increase, relative to the average mileage per piece registered pre-treatment (1.320 miles/piece), thereby supporting Hypothesis 3.

Table 15: Average effects of Amazon's integration on route density

	(1)		(2)	
	Estimate	Std. Err.	Estimate	Std. Err.
Integration × Treat (δ)	1.113***	(0.189)	1.173***	(0.247)
Observations	7,798		7,798	
R-squared	0.3357		0.3359	
Week FE	Yes		Yes	
Service area FE	Yes		Yes	
Control ($S_{i,t}$)	No		Yes	

Notes: 1. Robust standard errors are clustered at ZIP code level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specification in Column (2) includes as a predictor the weekly number of stops in each ZIP code area (S_{it}).

4 .1 Heterogeneity Effects

Table 16 reports the DID estimates for the treatment application effect on route density based on Equations (33) and (34). The estimates in Columns (1) to (4) account for the moderating effects of ZIP code area remoteness and proportion of fast deliveries separately using Equation (33). The estimates in Columns (5) and (6) account for these moderating effects simultaneously using Equation (34). Since the R^2 values obtained for the specifications in Columns (5) and (6) are higher than those for the specifications in the other columns, we will focus on these columns for the interpretation of the results.

The results show that an increase in remoteness of a ZIP code area amplifies the negative impact of Amazon's integration on route density, consistent with Hypothesis 4a. As we include ZIP code area fixed effects as part of our specification, this effect has a within-ZIP code area interpretation (Fitzmaurice *et al.*, 2012; Miller *et al.*, 2018). Specifically, every one-mile increase in our weekly measure of remoteness for each

ZIP code area amplifies by 0.069 miles Amazon's average vertical integration effect in the area's route density. This translates into a 0.859% $\left(0.069 \times \frac{16.430}{1.320}\right)$ decrease in density for every 1% increase in the weighted distance between each treatment ZIP code area and its hubs. [28] The proportion of fast deliveries also amplifies the impact of Amazon's integration, consistent with Hypothesis 4b. In particular, for every one percentage point increase in the proportion of fast deliveries in a ZIP code area, the mileage per piece in the area increases by 0.072. [29] This translates into a 5.45% (0.072/1.320) decrease in route density. [30]

Table 16: Heterogeneity effects of Amazon's integration on route density

	(1)	(2)	(3)	(4)	(5)	(6)
Integration × Treat (δ)	0.175	0.185	-0.519*	-0.461*	-1.707**	-1.695**
	(0.479)	(0.478)	(0.294)	(0.253)	(0.701)	(0.671)
Integration × Treat						
× Remoteness	0.058*	0.058*			0.069**	0.069**
	(0.034)	(0.034)			(0.031)	(0.031)
Integration × Treat						
× Proportion of Fast Deliveries			6.957***	6.985***	7.206***	7.210***
			(1.753)	(1.776)	(1.806)	(1.823)
Remoteness (Miles)	0.034	0.034			0.036	0.036
	(0.057)	(0.057)			(0.050)	(0.050)
Proportion of Fast Deliveries			-0.555	-0.592	-0.722	-0.727
			(0.623)	(0.634)	(0.616)	(0.631)
Observations	7,798	7,798	7,798	7,798	7,798	7,798
R-squared	0.343	0.343	0.406	0.407	0.416	0.416
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Service area FE	Yes	Yes	Yes	Yes	Yes	Yes
Control ($S_{i,t}$)	No	Yes	No	Yes	No	Yes

Notes: 1. Robust standard errors are clustered at ZIP code level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specifications in Columns (2), (4), and (6) include as a predictor the weekly number of stops in each ZIP code area (S_{it}).

[28] The average for this distance is 16.430 miles.

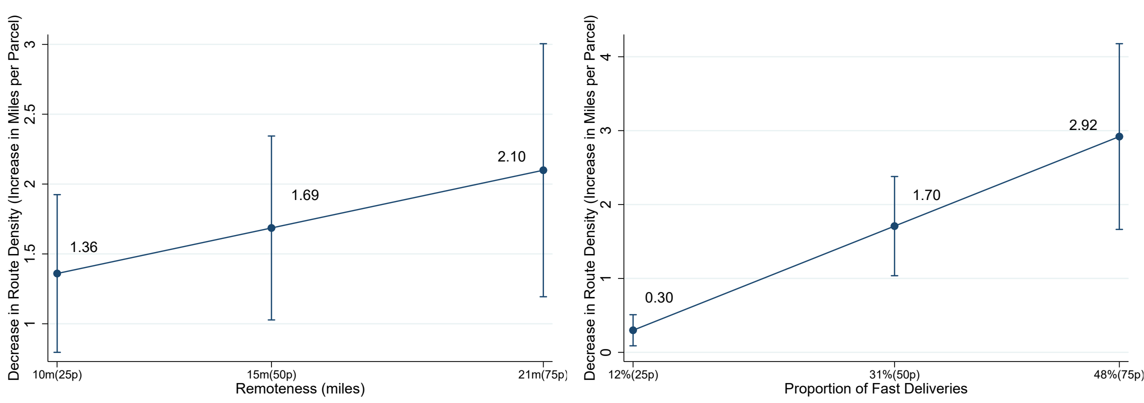
[29] Note that the proportion of fast deliveries is a percentage and therefore its coefficient is interpreted as a unit change in miles per piece with a 100 percentage point change in the percentage of same-day deliveries.

[30] These results are consistent with those obtained after logging remoteness and proportion of fast deliveries.

Having evaluated the average moderating effects by the level of remoteness and the proportion of fast deliveries, we next assess the magnitude of the decrease in route density evaluated at the interquantile range of these two moderators. This will allow us to understand which moderator is more detrimental to route density following Amazon's decision to vertically integrate. Figure 12 plots the marginal effect of Amazon's integration on the increase in mileage per parcel at the 25th, 50th, and 75th percentiles of each moderator while holding the other one equal to its mean value. As shown in the figure, the rate of increase in mileage per parcel is steeper across the different quartiles in the proportion of fast deliveries compared to the rate of increase across the quartiles in the level of remoteness. As the level of remoteness changes from the 25th percentile (10 miles) to the 75th percentile (21 miles), the increase in the mileage per parcel rises from 1.36 to 2.10, whereas the increase in the mileage per parcel rises from 0.3 miles per parcel to 2.92 miles per parcel as the proportion of fast deliveries changes from the 25th percentile (12%) to the 75th percentile (48%).

Figure 12: Decrease in route density varying with remoteness and proportion of fast deliveries

(A) Remoteness (B) Proportion of fast deliveries



Note: The graphs are generated based on estimates in Table 16 with a confidence interval at the 95% level. The marginal effects are evaluated at the interquartile values of each moderator while holding the other one equal to its mean. The x-axis labels display the value of each moderator at the 25th, 50th, and 75th percentiles. For example, at the 25th percentile, remoteness is 10 miles and the proportion of fast deliveries is 12%. All estimates are statistically significant at the 1% level.

4 .2 Robustness Tests

Although the approach we use to define the treatment and control groups has been well established to overcome selection biases, we re-estimated the values of the coefficients in Tables 14-16 while using alternative approaches based on two different propensity score procedures (propensity score weighting and propensity score matching). In both of these procedures, we calculate propensity scores that measure the similarities among different attributes between the ZIP code areas in the treatment and control groups. The goal is to maximize the similarity between the treatment and control groups conditional on the distribution of these attributes across the ZIP code areas in these groups (Rosenbaum and Rubin, 1983). The attributes we used for each ZIP code area include operational variables consisting of average miles per piece, remoteness, and proportion of fast deliveries, as well as demographic variables consisting of the number of households, the median annual household income, the number of retail establishments, and the number of accommodation and food service establishments. We obtained the measurements for the former two demographic attributes from the US Census Bureau American Community Survey (2017) and for the latter two from the US Census Bureau County Business Patterns (2016). We present the results obtained from the propensity score weighting and propensity score matching procedures in Appendix I and J, respectively. As shown in these appendices, the findings we obtained in this section are consistent with those obtained using both matching procedures.

As part of our robustness tests, we also consider the potential for false significance of our estimates by examining whether the observed reduction in route density caused by Amazon's integration occurred purely by chance. Therefore, we examine how the DID estimation performs on placebo events, where treated ZIP code areas and the week when Amazon's integration took place are chosen at random (Lonati *et al.*, 2018). We first draw a week at random from a uniform distribution among the weeks beginning on February 4, February 11, February 18, March 11, March 18, March 25, April 29, May 06, and May 13 and then code Amazon's integration placebo as 1 for the chosen week and all the following weeks. [31] Second, we select half of the ZIP code areas at random and designate them as the treatment group. We then estimate the DID parameter of the placebo events using Equation (32). Following Bertrand *et al.* (2004), we repeated this procedure 50 times, each time drawing a new placebo week of Amazon's integration and the treated ZIP code areas at random. We present the results in Figures 23a, 23b, and 23c of Appendix K, each corresponding to a result from the placebo tests based on the analyses using no propensity score procedures, using propensity score weighting, and using propensity score matching, respectively. We found that only 3 of the 150 placebo models across the figures produced a coefficient that is statistically different from zero at the 5% significance level.

5 Discussion

Our findings carry important implications for the operating costs in the transportation system we analyzed. Below, we quantify the impact on operating costs caused by decreases in route density after Amazon's integration and how this impact is amplified by the remoteness and the proportion of fast deliveries of a ZIP code area. We then expand on a variety of potential anti-competitive service and pricing outcomes stemming from these impacts on costs and put forth different public policy

[31] We chose to use placebo events in these weeks to ensure having enough observations before and after these events. As the integration in the ZIP code areas in the treatment groups occurred in three separate months (February, March, and May of 2017), the pool of weeks eligible for a placebo event for a particular ZIP code area is based on the month of the area's treatment date. For example, for the treatment groups with the treatment dates in February, the weeks eligible for placebo events are those beginning on February 4, February 11, and February 18 while for the treatment groups with treatment dates in March, the weeks eligible are those beginning on March 11, March 18, and March 25.

remedies that could be implemented to address these sources of anti-competitiveness in the last-mile delivery industry.

5 .1 Impacts on Operating Costs

For each ZIP code area, operating costs will increase as a result of reductions in route density, which is measured as $\Delta_y = \Delta_y^{integration} + \Delta_y^{RM} + \Delta_y^{FD}$. The first term is the average decrease in route density in a ZIP code area, Δ_y, following Amazon's integration without considering the moderating effects. The second and third terms correspond to the average decrease in route density amplified by the level of remoteness and the proportion of fast deliveries in the ZIP code area. Based on the estimates in Table 15, Δ_y is 1.173 miles per piece without considering the moderating effects. When the moderating effects are accounted for, the decrease in route density is measured as $\Delta_y = -1.695 + 0.069 \times RM + 7.210 \times FD$ (based on the estimates in Column 6 of Table 16).

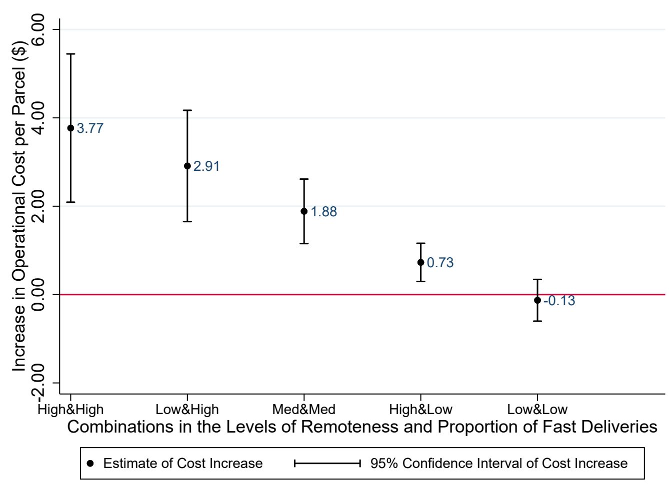
The increase in operating costs is a product of the decrease in route density and the operating cost per mile (i.e., $\Delta_c = \Delta_y \times$cost per mile). Based on a standard operating cost rate of \$0.535/mile per commercial vehicle and a labor rate of \$0.628/mile per driver, [32] we calculate the operating cost per mile as \$1.163 (\$0.535/mile + \$0.628/mile). Therefore, the increase in operating costs is $\Delta_c = \Delta_y \times \$1.163$.

On average, the increase in costs following the vertical integration of Amazon's platform across the ZIP code areas in our sample equals \$1.36 per piece ($\Delta_c = \Delta_y \times$ cost per mile = \$1.173 miles/piece $\times$ \$1.163/mile). As the increase in the operating costs depends on the values of remoteness and proportion of fast deliveries per ZIP

[32] The operating cost rate of \$0.535/mile is obtained from the Internal Revenue Service (2017). It includes depreciation, insurance, repairs, tires, maintenance, and gas. To obtain the labor rate of \$0.628/mile, we assume an average driving speed of 40 mph and an average driver salary of \$25/hour, consistent with compensation at Amazon (2020).

code area, we also evaluate Δ_c across different percentiles for these two variables. We considered percentiles at three different levels: high (75th percentile), medium (50th percentile), and low (25th percentile). For the high percentile level, the values for remoteness and proportion of fast deliveries are 21 miles and 48%, respectively. For the medium and low percentile levels, the values are: 15 miles and 31%, and 10 miles and 12%, respectively. Figure 13 presents the average estimations for Δ_c (including the 95 percent confidence intervals) across different percentile combinations for remoteness and proportion of fast deliveries. We also provide a summary of these results in Table 17 on a per-parcel and per-route basis (assuming that, on average, drivers deliver 30 parcels per route).

Figure 13: Cost increase by service quality



(1) All estimates are significantly different from 0 ($p < 0.01$), except for -0.13 (not significantly different from 0 at $p = 0.592$). (2) High, medium, and low levels of remoteness correspond to values at the 75th percentile (21 miles), 50th percentile (15 miles), and 25th percentile (10 miles), respectively. High, medium, and low levels in the proportion of fast deliveries correspond to values at the 75th percentile (48%), 50th percentile (31%), and 25th percentile (12%), respectively.

Table 17: Operating cost impact of Amazon's vertical integration

Combinations in the Levels of Remoteness and Proportion of Fast Deliveries	Increase in miles per parcel Δ_y	Total increase in operating cost (vehicular cost + labor cost)	
		per parcel Δ_c	per route
Average (without moderating effects)	1.173***	\$1.36	\$40.93
High&High	3.25***	\$3.77	\$113.10
Low&High	2.51***	\$2.91	\$87.30
Med&Med	1.62***	\$1.88	\$56.40
High&Low	0.63***	\$0.73	\$21.90
Low&Low	-0.11	\$-0.13	-\$3.90

*,** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.

5 .2 Potential Service Outcomes Caused by Increases in Costs Following Amazon's Integration

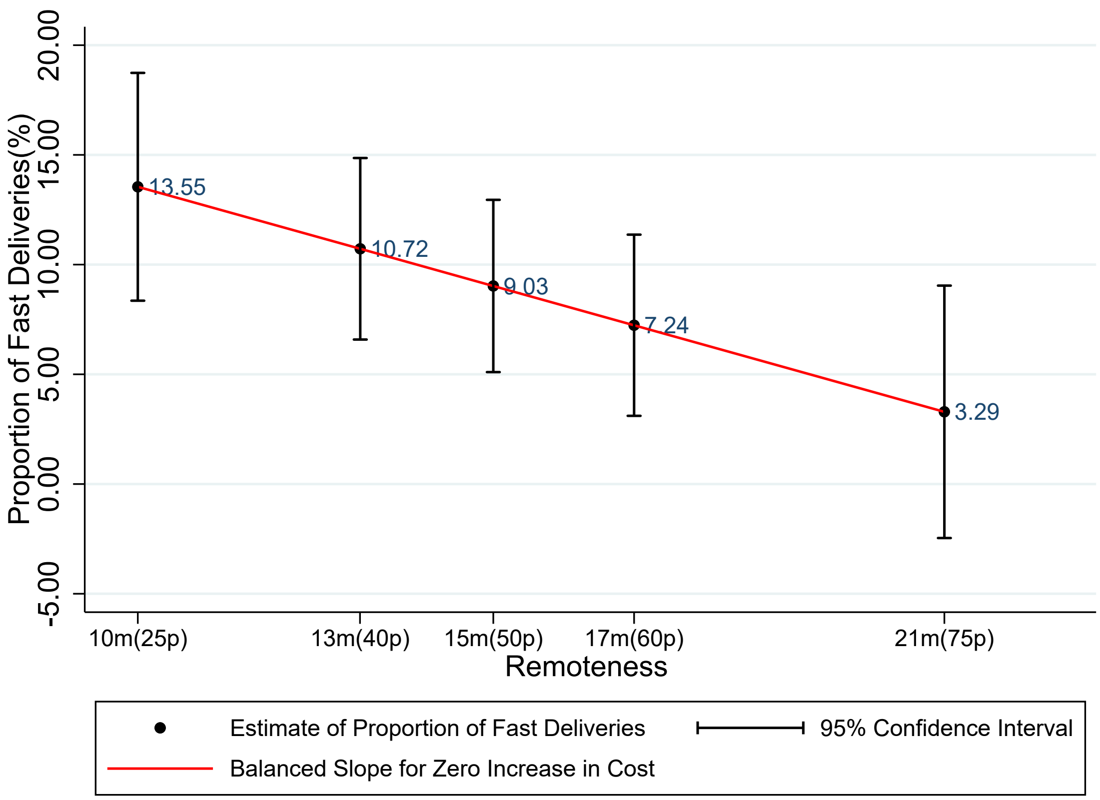
Based on Figure 13 and Table 17, we obtain the largest Δ_c value (\$3.77 per parcel, significantly different from 0) when using high values (75th percentile) for both variables (remoteness and proportion of fast deliveries) and the lowest Δ_c value (-\$0.13 per parcel, not significantly different from zero) when using low values (25th percentile) for both variables. These findings suggest that to mitigate detrimental cost effects caused by the reduction in route density following Amazon's integration, the last-mile delivery firm may be compelled to reduce its service levels in terms of fast delivery and reach to remote areas affected by this integration. This strategy is comparable to that involving reductions in store network density implemented by-brick and-mortar retailers after their introduction of superstore formats that allowed them to obtain scale economies in their store deliveries (Holmes, 2001). [33]

We are also able to compare the average value of Δ_c obtained when using a low percentile value for remoteness and a high percentile value for proportion of

[33]We thank an anonymous reviewer for pointing this out.

fast deliveries against the average Δ_c value obtained when using a high value for remoteness and a low value for proportion of fast deliveries. As shown in Figure 13, the average value for Δ_c is substantially larger for the former case than for the latter scenario (\$2.91 vs. \$0.73 per parcel). This suggests that, following Amazon's vertical integration, offering a high proportion of fast deliveries in affected ZIP code areas becomes more expensive than serving a higher proportion of affected remote ZIP code areas.

Figure 14: Iso-cost curve for zero cost increase



All proportions of fast delivery estimates are significantly different from 0 ($p < 0.01$), except for 3.29% (not significantly different from 0 at $p = 0.262$).

Figure 14 reinforces this point. It shows the percentile levels of remoteness and corresponding proportions of fast deliveries per ZIP code area that result in a zero cost increase (Δ_c), following Amazon's vertical integration. The figure shows that to prevent a cost increase across different percentile levels of remoteness, the proportion of fast deliveries per ZIP code area needs to be maintained at a low level (from 3.29% for a ZIP code area with a 75th percentile level of remoteness to 13.55% for a ZIP code area with a 25th percentile level of remoteness). Therefore, mitigating increases in costs following Amazon's vertical integration will require severely restricting the speed of service available across ZIP code areas. This, in turn, will make it more difficult for Amazon's retail rivals to rely on last-mile carriers in order to compete based on fast deliveries to consumers.

5 .3 Potential Pricing Outcomes Caused by Increases in Costs Following Amazon's Integration

The significant magnitudes we observed for the increases in costs caused by the integration of Amazon's platform also carry important pricing implications for the industry where our focal last-mile delivery firm competes. Consider the following scenarios that the firm faces before and after Amazon's platform integration.

- Pre-Integration: Let the average unit cost of fulfillment be denoted by c, the average number of parcels delivered for Amazon (resp., non-Amazon retailers) be N_A (resp., N_{NA}), and the average unit price charged to Amazon (resp., non-Amazon retailers) be r_A (resp., r_{NA}). The firm's profit is as follows:

$$\Pi_{pre} = \underbrace{N_A\,(r_A - c)}_{\text{Profit from Amazon}} + \underbrace{N_{NA}\,(r_{NA} - c)}_{\text{Profit from non-Amazon retailers}}.$$

- Post-Integration: The average cost of fulfillment increases by Δ_c; thus, $c \to c+\Delta_c$. The firm's profit is as follows:

$$\Pi_{post} = N_{NA}\,(r_{NA} - (c + \Delta_c)).$$

For fixed r_{NA}, the firm's profit decreases by $N_A(r_A - c) + N_{NA}\Delta_c$: The first term is the direct profit from Amazon, and the second term is the strategic externality that Amazon's vertical integration creates on non-Amazon deliveries. The firm has one of two solutions:

- Uniform Pricing (no price discrimination across retailers based on market share): If the firm follows a uniform pricing strategy, then, Amazon's exit leads to an increase in c. This eventually leads to an increase in r_{NA}. That is, every retailer is charged a higher price.

- If the firm is allowed to price differentiate retailers based on their volumes, by comparing the two profits, we have:

$$\underbrace{r_A - c}_{\text{Firm's unit margin from Amazon's deliveries}} \geq \underbrace{-\frac{N_{NA}}{N_A}\Delta_c}_{\text{Cross-Subsidization of Amazon's deliveries by non-Amazon retailers}}.$$

That is, the firm may incur a lower profit margin (or even a loss) on Amazon's deliveries, so as to encourage Amazon to participate. The maximum loss that the firm is willing to incur is the RHS: the product of relative market shares, $\frac{N_{NA}}{N_A}$, and the increase in the fulfillment cost due to Amazon's exit, Δ_c. Let c_A denote Amazon's own cost of fulfillment. For Amazon's participation to be rational, we require:

$$c_A \geq r_A,$$

Rearranging these inequalities, we get,

$$c_A \geq r_A \geq c - \frac{N_{NA}}{N_A}\Delta_c.$$

Therefore, the firm may charge Amazon a lower rate to incentivize participation, but increase the rate it charges to other retailers. That is, Amazon's deliveries will be cross-subsidized by other retailers.

5 .4 Public Policy Implications

Cross-subsidization can distort competition by giving a powerful platform like Amazon's an advantage, since Amazon will pay significantly lower rates than those paid by its retail rivals. Competitive distortions can also result from last-mile transportation firms' decisions to curtail their levels of service in order to limit increases in costs brought about by Amazon's vertical integration into their industry. Despite the evidence this paper presents about these potential distortions, existing antitrust policy has been mainly concerned with ensuring that prices in these systems remain low and outputs high, without getting involved in overseeing market power dynamics that may ultimately be responsible for these prices and outputs (Khan, 2016). It is possible that the economics of the last-mile transportation industry bear responsibility for these power dynamics, in which case there are at least two approaches public policy experts could consider. One approach is to promote industry governance through competition. This may involve the use of policing forms of vertical integration that dominant e-commerce platforms, like Amazon's, can use for anti-competitive purposes.

The other approach is to exploit the economic advantages that large platforms, including Amazon's, may bring to the table while regulating their ability to exploit their power. Our results show that Amazon's platform is sufficiently powerful to induce discriminatory pricing among last-mile transportation firms in these markets and to limit their ability to offer high service levels. Unlike monopolistic firms, Amazon may maintain the shipping and handling prices its platform charges to consumers low, obscuring its market power (Foer, 2014). However, Amazon can choose to operate its platform as a monopsony, whose buying power allows it to induce last-mile transportation firms to price their services discriminatorily in Amazon's favor. In this sense, Amazon may choose not to operate its platform as a monopolist and extract rents from consumers. Instead, it may choose to use its buying power to squeeze surplus from last-mile delivery firms participating in its platform (Dube *et al.*, 2020).

A platform like Amazon's presents a unique form of private power that manifests in its ability to set not just the terms of access to the platform itself but also the prices and returns for last-mile delivery firms and third-party sellers participating in the platform. While a monopolist's power is based on its control over the production and pricing of a particular good, a platform's power stems from its position as an intermediary controlling the relationships of the agents participating on its different sides (i.e., the last-mile delivery firms and third-party sellers). Once a platform reaches a critical mass, agents on one side of the paltform can become vulnerable to the platform's control over agents on the other side (Rahman, 2015).

How can public policy agencies monitor this phenomenon? Agencies have often relied on complaints filed by affected parties laying out antitrust cases emphasizing harm to consumers. At the end of 2019, a third-party seller participating in Amazon's platform filed such a complaint. In it, the seller accused Amazon of forcing it to use Amazon's in-house delivery services even though the rates for these services exceeded by as much as 35% of those available through outside last-mile delivery firms. This, in turn, forced the seller to increase the prices of the products it sells on the platform by 12%. According to the complaint, Amazon pushes sellers to use its services because doing so will make it more likely that their items will be listed prominently in the search results on Amazon's site. Moreover, third-party sellers who choose to use Amazon's services have the advantage of never being penalized by Amazon for delivery errors involving their products. This is not the case for sellers who choose to use outside services. These sellers are exposed to penalties, including being kicked out of the platform, for incurring delivery mistakes (Soper, 2019).

This particular complaint garnered significant attention because it laid out a case that emphasized harm to consumers – the traditional basis for antitrust cases in the U.S. Without any evidence of consumer harm, Amazon's use of its platform's market power to induce last-mile transportation firms to price their services discriminatorily in its favor or cut down their levels of service may not draw the same amount of attention unless it can be shown that such effects on last-mile transportation firms erode competition and consumer welfare (Hochstadt *et al.*, 2020).

What remedies are available to ameliorate outcomes harmful to consumer welfare and competition? Based on theoretical models, Qin *et al.* (2020, 2021) show that in-house delivery services by a platform like Amazon's can maximize consumer surplus when the orders being delivered are sold directly by the platform owner (i.e., Amazon). Third-party sellers' use of in-house delivery services to fulfill their orders is not conducive to maximizing consumer surplus because it does not guarantee a high level of service quality to consumers. Lai *et al.* (2018) also found through a theoretical model that by inducing third-party sellers to use its services available through the Fulfillment by Amazon (FBA) program, Amazon can blunt the sellers' ability to compete with Amazon on prices. In addition, according to Sun *et al.* (2020), the use of these services by third-party sellers increases their exposure to product returns by consumers participating on the platform. It is possible that these phenomena will also apply to Amazon's in-house delivery services.

Regulating Amazon's ability to exploit its platform's power may involve applying an "essential facilities doctrine" whereby last-mile transportation firms are granted access to delivering parcels from third-party sellers participating in the platform. This doctrine has been previously applied to ensuring access to networks in the communication and transportation industries (Lao, 2009; Ducci, 2020). Therefore, considering efforts by retailers like Amazon to leverage their platforms to dominate the last-mile transportation industry, it may make sense to apply this doctrine to ensure last-mile transportation firms have access to delivering the parcels from third-party sellers participating in these platforms. This access becomes particularly relevant for the delivery of parcels in areas that are expensive to serve, either because they are difficult to access or because they require faster deliveries.

Future research can shed further light on these issues. For instance, it is yet to be determined how the impact we observed following the vertical integration by Amazon's platform compares to that caused by the integration of other retailers' platforms, such as Target's. It is also unclear how this impact will hold in the long run. Because this form of vertical integration is a relatively recent phenomenon there is limited data available to estimate long-term effects. Furthermore, limited data access prevents us from evaluating the effects caused by Amazon's decision to vertically integrate its platform across more last-mile delivery firms. As more data on this phenomenon becomes available, future research may assess our results in relation to the effects on the operations of other last-mile delivery firms caused by decisions by Amazon, as well as other retailers, to vertically integrate their platforms.

In addition, research could empirically explore alternative motivations behind these platforms' entry into the last-mile transportation industry. For example, Borsenberger *et al.* (2018) provide analytical evidence that these platforms are more likely to deploy their own delivery networks in areas that are less costly to serve while leaving other areas open to last-mile delivery firms. They also show that integration focused solely on less expensive areas is more likely to have a negative effect on welfare than full integration. Empirical research could examine whether in fact these motivations are present among platforms operated by retailers such as Amazon. It is possible that these platforms may be able to integrate into more expensive areas only if they are able to include a markup in the delivery rates in these areas.

Finally, research could evaluate how various policy remedies could mitigate potential negative externalities of vertical integration. The increases in the mileage per parcel delivered that we observe by our study's last-mile transportation firm following Amazon's vertical integration may lead to greater greenhouse gas (GHG) emissions in the areas where these deliveries take place. Who bears responsibility for these externalities? The literature offers no clear answers. Obviously, these environmental costs would not arise but for Amazon's decision to vertically integrate its platform. Therefore, one could argue that Amazon should bear responsibility for these costs. At the same time, however, one could assign responsibility for these costs to existing route operators, such as the last-mile transportation firm in our study, since these costs are generated by the GHG emissions from the operators' own routes. Another issue is how to regulate these responsibilities. Imposing emission penalties on the parties responsible for environmental costs may not be effective from a system-wide perspective because these penalties can depress output below optimum social welfare levels (Sim et al. 2019). Instead, regulators may consider using incentive programs that subsidize firms' efforts to reduce environmental costs (Drake *et al.*, 2016). For instance, according to operators in the last-mile transportation industry, these programs have been very effective in reducing the upfront capital costs incurred in replacing gas/diesel vehicles with electric ones in their fleets. They have also contributed to offset part of the costs of ownership of electric vehicles (Leung and Peace, 2020). Clearly additional research is necessary to address these concerns.

6 Conclusion

In this paper, we examine how Amazon's decision to vertically integrate its retail platform and last-mile delivery operations can lead to anti-competitive outcomes as a result of a deterioration in the operating efficiency in the routes served by a last-mile delivery firm. Through our analysis, we provide an initial benchmark of the effects on efficiency in the last-mile delivery industry caused by the vertical integration of platforms owned by Amazon and other retailers. Moreover, we provide a basis to highlight possible anti-competitive implications for the industry caused by these platforms' decisions to vertically integrate. In so doing, we help inform the ongoing debate surrounding antitrust concerns about these platforms and their market power. [34] Finally, we identify opportunities for future research, including the investigation of the potential rise of negative externalities following these platforms' vertical integration.

[34] Some of these concerns were recently aired in Congress. Please refer to: Hearings, Online Platforms and Market Power, Part 6: Examining the Dominance of Amazon, Apple, Facebook, and Google, U.S. House Judiciary Committee, Subcommittee on Antitrust, Commercial, and Administrative Law (July 29, 2020).

2020 USDA Local Food Promotion Program URL https://www.ams.usda.gov/sites/default/files/media/LFPPDescriptionofFundedProjects2020.pdf, accessed on May 20, 2021. (2020).

Abadie, A., "Semiparametric difference-in-differences estimators", The Review of Economic Studies **72**, 1, 1–19 (2005).

Agatz, N., A. Campbell, M. Fleischmann and M. Savelsbergh, "Time slot management in attended home delivery", Transportation Science **45**, 3, 435–449 (2011).

Agatz, N., A. M. Campbell, M. Fleischmann, J. Van Nunen and M. Savelsbergh, "Revenue management opportunities for internet retailers", Journal of Revenue and Pricing Management **12**, 2, 128–138 (2013).

Ahumada, O. and J. R. Villalobos, "Application of planning models in the agri-food supply chain: A review", European journal of Operational research **196**, 1, 1–20 (2009).

Ailawadi, K. L., E. T. Bradlow, M. Draganska, V. Nijs, R. P. Rooderkerk, K. Sudhir, K. C. Wilbur and J. Zhang, "Empirical models of manufacturer-retailer interaction: A review and agenda for future research", Marketing Letters **21**, 3, 273–285 (2010).

Ailawadi, K. L., K. Gedenk, C. Lutzky and S. A. Neslin, "Decomposition of the sales impact of promotion-induced stockpiling", Journal of Marketing Research **44**, 3, 450–467 (2007).

Akşin, Z., B. Ata, S. M. Emadi and C.-L. Su, "Structural estimation of callers' delay sensitivity in call centers", Management Science **59**, 12, 2727–2746 (2013).

Alimonti, R., L. Mautino and L. Stammati, "E-commerce growth: Competition and regulatory implications for the postal sector", in "The Changing Postal Environment", pp. 167–181 (Springer, 2020).

Allon, G., M. Cohen and P. Sinchaisri, "The impact of behavioral and economic drivers on gig economy workers", URL https://ssrn.com/abstract=3274628 (2018).

Amazon, "Start earning", URL https://flex.amazon.com, accessed on July 30, 2020. (2020).

Angrist, J. D. and J.-S. Pischke, *Mostly harmless econometrics: An empiricist's companion* (Princeton University Press, 2008).

Arcidiacono, P. and P. B. Ellickson, "Practical methods for estimation of dynamic discrete choice models", Annu. Rev. Econ. **3**, 1, 363–394 (2011).

Arcidiacono, P. and J. B. Jones, "Finite mixture distributions, sequential likelihood and the EM algorithm", Econometrica **71**, 3, 933–946 (2003).

Arcidiacono, P. and R. A. Miller, "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity", Econometrica **79**, 6, 1823–1867 (2011).

Armstrong, M., "Competition in two-sided markets", The RAND Journal of Economics **37**, 3, 668–691 (2006).

Arnosti, N., R. Johari and Y. Kanoria, "Managing congestion in matching markets", Manufacturing & Service Operations Management **23**, 3, 620–636 (2021).

Ashenfelter, O. and D. Card, *Handbook of labor economics* (Elsevier, 2010).

Babar, Y. and G. Burtch, "Examining the heterogeneous impact of ride-hailing services on public transit use", Information Systems Research **31**, 3, 820–834 (2020).

Bai, J., K. C. So, C. S. Tang, X. Chen and H. Wang, "Coordinating supply and demand on an on-demand service platform with impatient customers", Manufacturing&Service Operations Management (2018).

Bamberger, K. A. and O. Lobel, "Platform market power", Berkeley Technology Law Journal **32**, 1051 (2017).

Bell, D. R., T.-H. Ho and C. S. Tang, "Determining where to shop: Fixed and variable costs of shopping", Journal of Marketing Research pp. 352–369 (1998).

Benjaafar, S., J.-Y. Ding, G. Kong and T. Taylor, "Labor welfare in on-demand service platforms", Manufacturing & Service Operations Management **forthcoming** (2020).

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

Berman, J., "2020 parcel express roundtable: You get what you negotiate", Logistics Management URL `https://www.logisticsmgmt.com/article/2020_parcel_express_roundtable_you_get_what_you_negotiate`, accessed on July 30, 2020. (2020).

Berry, S., J. Levinsohn and A. Pakes, "Automobile prices in market equilibrium", Econometrica: Journal of the Econometric Society pp. 841–890 (1995).

Berti, G., C. Mulligan and H. Yap, "Digital food hubs as disruptive business models based on coopetition and "shared value" for sustainability in the agri-food sector", in "Global Opportunities for Entrepreneurial Growth: Coopetition and Knowledge Dynamics within and across Firms", (Emerald Publishing Limited, 2017).

Bertrand, M., E. Duflo and S. Mullainathan, "How much should we trust differences-in-differences estimates?", The Quarterly Journal of Economics **119**, 1, 249–275 (2004).

Besanko, D., S. Gupta and D. Jain, "Logit demand estimation under competitive pricing behavior: An equilibrium framework", Management Science **44**, 11-part-1, 1533–1547 (1998).

Bhargava, H. K., B. C. Kim and D. Sun, "Commercialization of platform technologies: Launch timing and versioning strategy", Production and Operations Management **22**, 6, 1374–1388 (2013).

Bielaczyc, N., R. Pirog, J. Fisk, J. Fast and P. Sanders, "Findings of the 2019 national food hub survey", URL `https://www.canr.msu.edu/resources/2019-food-hub-survey`, accessed on May 20, 2021. (2020).

Bimpikis, K., W. J. Elmaghraby, K. Moon and W. Zhang, "Managing market thickness in online business-to-business markets", Management Science **66**, 12, 5783–5822 (2020).

Bloom, J. D. and C. C. Hinrichs, "The long reach of lean retailing: Firm embeddedness and wal-mart's implementation of local produce sourcing in the us", Environment and Planning A: Economy and Space **49**, 1, 168–185 (2017).

Bond, J. K., D. Thilmany and C. Bond, "What influences consumer choice of fresh produce purchase location?", Journal of Agricultural and Applied Economics **41**, 1, 61–74 (2009).

Bonnet, C., P. Dubois, S. B. Villas Boas and D. Klapper, "Empirical evidence on the role of nonlinear wholesale pricing and vertical restraints on cost pass-through", Review of Economics and Statistics **95**, 2, 500–515 (2013).

Borsenberger, C., H. Cremer, D. Joram and J.-M. Lozachmeur, "Vertical integration in the e-commerce sector", in "New Business and Regulatory Strategies in the Postal Sector", pp. 143–160 (Springer International Publishing, 2018).

Boushey, H. and J. S. Glynn, "There are significant business costs to replacing employees", URL `https://www.americanprogress.org/wp-content/uploads/2012/11/CostofTurnover.pdf`, accessed on August 24, 2020. (2012).

Braeutigam, R. R., "Learning about transport costs", Essays in transportation economics and policy: A handbook in honor of John R. Meyer **3**, 57–97 (1999).

Briesch, R. A., P. K. Chintagunta and E. J. Fox, "How does assortment affect grocery store choice?", Journal of Marketing research **46**, 2, 176–189 (2009).

Bucklin, R. E., S. Gupta and S. Siddarth, "Determining segmentation in sales response across consumer purchase behaviors", Journal of Marketing Research pp. 189–197 (1998).

Burtch, G., S. Carnahan and B. N. Greenwood, "Can you gig it? an empirical examination of the gig economy and entrepreneurial activity", Management Science **64**, 12, 5497–5520 (2018).

Cachon, G. P., "Supply chain coordination with contracts", Handbooks in operations research and management science **11**, 227–339 (2003).

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

Cachon, G. P., K. M. Daniels and R. Lobel, “The role of surge pricing on a service platform with self-scheduling capacity”, in “Sharing Economy”, pp. 101–113 (Springer, 2019).

Caillaud, B. and B. Jullien, “Chicken & egg: Competition among intermediation service providers”, RAND journal of Economics pp. 309–328 (2003).

Cameron, A. C. and P. K. Trivedi, “Regression-based tests for overdispersion in the Poisson model”, Journal of econometrics **46**, 3, 347–364 (1990).

Cao, J., M. Olvera-Cravioto and Z.-J. Shen, “Last-mile shared delivery: A discrete sequential packing approach”, URL `http://arxiv.org/abs/1805.05012` (2018).

Carpio, C. E. and O. Isengildina-Massa, “Consumer willingness to pay for locally grown products: The case of South Carolina”, Agribusiness: An International Journal **25**, 3, 412–426 (2009).

Castillo, V. E., J. E. Bell, W. J. Rose and A. M. Rodrigues, “Crowdsourcing last mile delivery: Strategic implications and future research directions”, Journal of Business Logistics **39**, 1, 7–25 (2018).

Cattani, K., O. Perdikaki and A. Marucheck, “The perishability of online grocers”, Decision Sciences **38**, 2, 329–355 (2007).

Chan, K. W., S. Y. Li, J. Ni and J. J. Zhu, “What feedback matters? the role of experience in motivating crowdsourcing innovation”, Production and Operations Management (2020).

Chen, M. K., “Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform”, in “Proceedings of the 2016 ACM Conference on Economics and Computation”, pp. 455–455 (2016).

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

Chintagunta, P. K., “Investigating category pricing behavior at a retail chain”, Journal of Marketing Research **39**, 2, 141–154 (2002).

Chu, J. and P. Manchanda, “Quantifying cross and direct network effects in online consumer-to-consumer platforms”, Marketing Science **35**, 6, 870–893 (2016).

Chung, D. J., T. Steenburgh and K. Sudhir, “Do bonuses enhance sales productivity? a dynamic structural analysis of bonus-based compensation plans”, Marketing Science **33**, 2, 165–187 (2014).

Cox, D. R., "Regression models and life-tables", Journal of the Royal Statistical Society: Series B (Methodological) **34**, 2, 187–202 (1972).

Crémer, J., Y.-A. de Montjoye and H. Schweitzer, "Competition policy for the digital era", Report for the European Commission (2019).

Cullen, Z. and C. Farronato, "Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms", Management Science (2020).

Darby, K., M. T. Batte, S. Ernst and B. Roe, "Decomposing local: A conjoint analysis of locally produced foods", American Journal of Agricultural Economics **90**, 2, 476–486 (2008).

De Croon, E. M., J. K. Sluiter, R. W. Blonk, J. P. Broersen and M. H. Frings-Dresen, "Stressful work, psychological job strain, and turnover: a 2-year prospective cohort study of truck drivers.", Journal of applied psychology **89**, 3, 442 (2004).

De Groote, O. and F. Verboven, "Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems", American Economic Review **109**, 6, 2137–72 (2019).

DellaVigna, S. and M. Gentzkow, "Uniform pricing in us retail chains", The Quarterly Journal of Economics **134**, 4, 2011–2084 (2019).

Deng, Q., X. Fang and Y. F. Lim, "Urban consolidation center or peer-to-peer platform? the solution to urban last-mile delivery", Production and Operations Management **30**, 4, 997–1013 (2021).

Dhanorkar, S., "Environmental benefits of Internet-enabled C2C closed-loop supply chains: A quasi-experimental study of Craigslist", Management Science **65**, 2, 660–680 (2018).

Diamond, R., T. McQuade and F. Qian, "The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco", American Economic Review **109**, 9, 3365–94 (2019).

Dinerstein, M., L. Einav, J. Levin and N. Sundaresan, "Consumer price search and platform design in internet commerce", American Economic Review **108**, 7, 1820–59 (2018).

Dobson, P. W. and R. Inderst, "The waterbed effect: Where buying and selling power come together", Wisconsin Law Review p. 331 (2008).

Draganska, M. and D. C. Jain, "Product-line length as a competitive tool", Journal of Economics & Management Strategy **14**, 1, 1–28 (2005).

Drake, D. F., P. R. Kleindorfer and L. N. Van Wassenhove, "Technology choice and capacity portfolios under emissions regulation", Production and Operations Management **25**, 6, 1006–1025 (2016).

Dube, A., J. Jacobs, S. Naidu and S. Suri, “Monopsony in online labor markets”, American Economic Review: Insights **2**, 1, 33–46 (2020).

Ducci, F., *Natural Monopolies in Digital Platform Markets* (Cambridge University Press, 2020).

Duggan, W., “Latest e-commerce market share numbers highlight amazon’s dominance”, Yahoo Finance URL `https://finance.yahoo.com/news/latest-e-commerce-market-share-185120510.html`, accessed on July 30, 2020. (2020).

Eisenmann, T., G. Parker and M. W. Van Alstyne, “Strategies for two-sided markets”, Harvard business review **84**, 10, 92 (2006).

Ellickson, P. B., S. Misra and H. S. Nair, “Repositioning dynamics and pricing strategy”, Journal of Marketing Research **49**, 6, 750–772 (2012).

Emadi, S. M. and B. R. Staats, “A structural estimation approach to study agent attrition”, Management Science **66**, 9, 4071–4095 (2020).

Evans, D. S. and R. Schmalensee, “Failure to launch: Critical mass in platform businesses”, Review of Network Economics **9**, 4 (2010).

Evans, D. S. and R. Schmalensee, *Matchmakers: The new economics of multisided platforms* (Harvard Business Review Press, 2016).

Fadlon, I. and T. H. Nielsen, “Family health behaviors”, American Economic Review **109**, 9, 3162–91 (2019).

Fang, H. and Y. Wang, “Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions”, International Economic Review **56**, 2, 565–596 (2015).

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

Fatehi, S. and M. R. Wagner, “Crowdsourcing last-mile deliveries”, (2021).

Fitzmaurice, G. M., N. M. Laird and J. H. Ware, *Applied longitudinal analysis*, vol. 998 (John Wiley & Sons, 2012).

Foer, F., “Amazon must be stopped”, New Republic URL `https://newrepublic.com/article/119769/amazons-monopoly-must-be-broken-radical-plan-tech-giant` (2014).

Fradkin, A., “Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb”, URL `https://ssrn.com/abstract=2939084` (2017).

Fraiberger, S. P. and A. Sundararajan, “Peer-to-peer rental markets in the sharing economy”, URL `https://conference.nber.org/conferences/2015/EoDs15/FraibergerSundararajanNBERDigitization0306.pdf` (2015).

Gardner, T. M., P. M. Wright and L. M. Moynihan, “The impact of motivation, empowerment, and skill-enhancing practices on aggregate voluntary turnover: The mediating effect of collective affective commitment”, Personnel psychology **64**, 2, 315–350 (2011).

Greene, W. H., *Econometric analysis* (Pearson Education India, 2003).

Gruber, J., “The incidence of mandated maternity benefits”, American Economic Review pp. 622–641 (1994).

Guda, H. and U. Subramanian, “Your uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives”, Management Science **65**, 5, 1995–2014 (2019).

Gunders, D. and J. Bloom, *Wasted: How America is losing up to 40 percent of its food from farm to fork to landfill* (Natural Resources Defense Council New York, 2017).

Gurvich, I., M. Lariviere and A. Moreno, “Operations in the on-demand economy: Staffing services with self-scheduling capacity”, in “Sharing Economy”, pp. 249–278 (Springer, 2019).

Hagiu, A., “Pricing and commitment by two-sided platforms”, The RAND Journal of Economics **37**, 3, 720–737 (2006).

Hagiu, A., “Two-sided platforms: Product variety and pricing structures”, Journal of Economics & Management Strategy **18**, 4, 1011–1043 (2009).

Halaburda, H., M. Jan Piskorski and P. Yıldırım, “Competing by restricting choice: The case of matching platforms”, Management Science **64**, 8, 3574–3594 (2017).

Hall, A. R., *Generalized method of moments* (Oxford university press, 2005).

Han, B. R., T. Sun, L. Y. Chu and L. Wu, “Connecting customers and merchants offline: Experimental evidence from the commercialization of last-mile stations at Alibaba”, URL `https://ssrn.com/abstract=3452769` (2019).

Hansen, L. P. and K. J. Singleton, “Generalized instrumental variables estimation of nonlinear rational expectations models”, Econometrica: Journal of the Econometric Society pp. 1269–1286 (1982).

Hardesty, S. D., “The growing role of local food markets”, American Journal of Agricultural Economics **90**, 5, 1289–1295 (2008).

Hausman, J., G. Leonard and J. D. Zona, “Competitive analysis with differenciated products”, Annales d’Economie et de Statistique pp. 159–180 (1994).

Heckman, J. J. and G. Sedlacek, “Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market”, Journal of Political Economy **93**, 6, 1077–1125 (1985).

Hesterman, O. B. and D. Horan, “The demand for “local” food is growing — here’s why investors should pay attention”, URL https://www.businessinsider.com/the-demand-for-local-food-is-growing-2017-4, accessed on February 13, 2020. (2017).

Ho, T.-H., N. Lim, S. Reza and X. Xia, “Om forum—causal inference models in operations management”, Manufacturing & Service Operations Management **19**, 4, 509–525 (2017).

Hochstadt, E., Y. Buchweitz and E. A. Rivas, “Platform conduct: Navigating new grounds”, The Antitrust Source URL https://www.weil.com/~/media/files/pdfs/2020/weil_antitrust_latestthinking_feb2020.pdf (2020).

Hoffman, M. and S. V. Burks, “Worker overconfidence: Field evidence and implications for employee turnover and firm profits”, Quantitative Economics **11**, 1, 315–348 (2020).

Holmes, T. J., “Bar codes lead to frequent deliveries and superstores”, RAND Journal of Economics pp. 708–725 (2001).

Hom, P. W., T. W. Lee, J. D. Shaw and J. P. Hausknecht, “One hundred years of employee turnover theory and research.”, Journal of applied psychology **102**, 3, 530 (2017).

Horst, M., E. Ringstrom, S. Tyman, M. Ward, V. Werner and B. Born, “Toward a more expansive understanding of food hubs”, Journal of Agriculture, Food Systems, and Community Development **2**, 1, 209–225 (2011).

Hotz, V. J. and R. A. Miller, “Conditional choice probabilities and the estimation of dynamic models”, The Review of Economic Studies **60**, 3, 497–529 (1993).

Huang, L. and M. D. Smith, “The dynamic efficiency costs of common-pool resource exploitation”, American Economic Review **104**, 12, 4071–4103 (2014).

Hwang, M. and S. Park, “The impact of Walmart supercenter conversion on consumer shopping behavior”, Management Science **62**, 3, 817–828 (2016).

Imbens, G. and J. Wooldridge, “Control function and related methods. what’s new in econometrics”, Lecture Notes **6** (2007).

Internal Revenue Service, “2017 standard mileage rates for business, medical and moving announced”, URL https://www.irs.gov/newsroom/2017-standard-mileage-rates-for-business-medical-and-moving-announced, accessed on July 30, 2020. (2017).

Iyengar, S. S. and M. R. Lepper, “When choice is demotivating: Can one desire too much of a good thing?”, Journal of personality and social psychology **79**, 6, 995 (2000).

Jansen, C., “What Amazon’s last-mile delivery ambitions mean for carriers”, Retail Dive URL https://www.retaildive.com/news/what-amazons-last-mile-delivery-ambitions-mean-for-carriers/547143/, accessed on June 14, 2020. (2019).

Jiang, B., K. Jerath and K. Srinivasan, “Firm strategies in the “mid tail” of platform-based retailing”, Marketing Science **30**, 5, 757–775 (2011).

Jiang, B. and L. Tian, “Collaborative consumption: Strategic and economic implications of product sharing”, Management Science **64**, 3, 1171–1188 (2016).

Jovanovic, B., “Job matching and the theory of turnover”, Journal of Political Economy **87**, 5, Part 1, 972–990 (1979).

Kabra, A., E. Belavina and K. Girotra, “Designing promotions to scale marketplaces”, Tech. rep., Working paper, INSEAD, Fontainebleau, France, URL http://akabra.com/documents/TaxiMarketplaces-06-11-2018.pdf (2016).

Kaiser, U. and J. Wright, “Price structure in two-sided markets: Evidence from the magazine industry”, International Journal of Industrial Organization **24**, 1, 1–28 (2006).

Kalouptsidi, M., P. T. Scott and E. Souza-Rodrigues, “Linear iv regression estimators for structural dynamic discrete choice models”, Journal of Econometrics **222**, 1, 778–804 (2021).

Kang, A., R. Lowery and M. Wardlaw, “The costs of closing failed banks: A structural estimation of regulatory incentives”, The Review of Financial Studies **28**, 4, 1060–1102 (2015).

Keane, M. P. and K. I. Wolpin, “The career decisions of young men”, Journal of Political Economy **105**, 3, 473–522 (1997).

Kenney, M. and J. Zysman, “The platform economy: restructuring the space of capitalist accumulation”, Cambridge journal of regions, economy and society **13**, 1, 55–76 (2020).

Kessler, F. and R. H. Stern, “Competition, contract, and vertical integration”, The Yale Law Journal **69**, 1, 1–129 (1959).

Keyes, D., “E-Commerce will make up 17% of all US retail sales by 2022 – and one company is the main reason - Business Insider”, URL https://www.businessinsider.com/e-commerce-retail-sales-2022-amazon-2017-8, accessed on February 22, 2020. (2017).

Khan, L. M., “Amazon’s antitrust paradox”, Yale Law Journal **126**, 710 (2016).

Khan, L. M., “The separation of platforms and commerce”, Columbia Law Review **119**, 4, 973–1098 (2019).

Kuksov, D. and J. M. Villas-Boas, “When more alternatives lead to less choice”, Marketing Science **29**, 3, 507–524 (2010).

Lai, G., H. Liu and W. Xiao, “Fulfilled by Amazon: A strategic perspective of competition at the E-commerce platform”, Available at SSRN 3270958 URL https://ssrn.com/abstract=3270958 (2018).

Lao, M., “Networks, access, and essential facilities: From terminal railroad to microsoft”, SMU Law Review **62**, 557 (2009).

Laseter, T., A. Tipping and F. Duiven, “The rise of the last-mile exchange”, URL https://www.strategy-business.com/article/The-Rise-of-the-Last-Mile-Exchange, accessed on July 30, 2020. (2018).

Leung, J. and J. Peace, “Insights on electric trucks for retailers and trucking companies”, URL https://www.c2es.org/site/assets/uploads/2020/02/Insights-On-Electric-Trucks-For-Retailers-And-Trucking-Companies.pdf, accessed on July 30, 2020. (2020).

Leyland, A., “Tesco’s major reset this autumn will have a huge impact on the market”, URL https://www.thegrocer.co.uk/leader/tescos-major-reset-will-have-a-huge-impact-on-the-market/512869.article, accessed on February 13, 2020. (2015).

Li, J., N. Granados and S. Netessine, “Are consumers strategic? structural estimation from the air-travel industry”, Management Science **60**, 9, 2114–2137 (2014).

Li, J. and S. Netessine, “Higher market thickness reduces matching rate in online platforms: Evidence from a quasiexperiment”, Management Science **66**, 1, 271–289 (2020).

Lian, Z. and G. Van Ryzin, “Optimal growth in two-sided markets”, Management Science (2021).

Lim, S. F. W., E. Rabinovich, S. Park and M. Hwang, “Shopping activity at warehouse club stores and its competitive and network density implications”, Production and Operations Management **30**, 1, 28–46 (2021).

Liu, X., Y. Cui and L. Chen, “Bonus competition in the gig economy”, URL https://ssrn.com/abstract=3392700 (2019).

Lonati, S., B. F. Quiroga, C. Zehnder and J. Antonakis, “On doing relevant and rigorous experiments: Review and recommendations”, Journal of Operations Management **64**, 19–40 (2018).

Low, S. A. and S. J. Vogel, “Direct and intermediated marketing of local foods in the United States”, Tech. Rep. 128, USDA-ERS Economic Research Report, URL https://papers.ssrn.com/abstract=2114361, accessed on February 13, 2020. (2011).

Martinez, S., *Local food systems; concepts, impacts, and issues* (Diane Publishing, 2010).

Matson, J., M. Sullins and C. Cook, "The role of food hubs in local food marketing", Tech. rep., United States Department of Agriculture, URL `https://ageconsearch.umn.edu/record/280771/files/sr73.pdf`, accessed on February 13, 2020. (2013).

McFadden, D., "Conditional logit analysis of qualitative choice behavior", pp. 105–142 (Academic Press, New York, 1974).

Mehta, N., S. Rajiv and K. Srinivasan, "Price uncertainty and consumer search: A structural model of consideration set formation", Marketing science **22**, 1, 58–84 (2003).

Mena, C., B. Adenso-Diaz and O. Yurt, "The causes of food waste in the supplier–retailer interface: Evidences from the uk and spain", Resources, Conservation and Recycling **55**, 6, 648–658 (2011).

Miller, J. W., D. C. Ganster and S. E. Griffis, "Leveraging big data to develop supply chain management theory: The case of panel data", Journal of Business Logistics **39**, 3, 182–202 (2018).

Mims, C., "In a tight labor market, gig workers get harder to please", URL `https://www.wsj.com/articles/in-a-tight-labor-market-gig-workers-get-harder-to-please-11556942404`, accessed on February 22, 2020. (2019).

Mincer, J., *Schooling, Experience, and Earnings* (New York: Columbia University Press, 1974).

Misra, S. and H. S. Nair, "A structural model of sales-force compensation dynamics: Estimation and field implementation", Quantitative Marketing and Economics **9**, 3, 211–257 (2011).

Mitchell, T. R. and L. R. James, "Building better theory: Time and the specification of when things happen", Academy of Management Review **26**, 4, 530–547 (2001).

Moon, K., P. Bergemann, D. Brown, A. Chen, J. Chu, E. Eisen, G. Fischer, P. K. Loyalka, S. Rho and J. Cohen, "Manufacturing productivity with worker turnover", URL `https://ssrn.com/abstract=3248075` (2018).

Mukhopadhyay, T., S. Rajiv and K. Srinivasan, "Information technology impact on process output and quality", Management Science **43**, 12, 1645–1659 (1997).

Murphy, A., "A dynamic model of housing supply", American Economic Journal: Economic Policy **10**, 4, 243–267 (2018).

Musalem, A., M. Olivares, E. T. Bradlow, C. Terwiesch and D. Corsten, "Structural estimation of the effect of out-of-stocks", Management Science **56**, 7, 1180–1197 (2010).

Nair, H., P. Chintagunta and J.-P. Dubé, "Empirical analysis of indirect network effects in the market for personal digital assistants", Quantitative Marketing and Economics **2**, 1, 23–58 (2004).

Nakamura, E. and D. Zerom, "Accounting for incomplete pass-through", The Review of Economic Studies **77**, 3, 1192–1230 (2010).

Nelson, P., "Information and consumer behavior", Journal of political economy **78**, 2, 311–329 (1970).

Nikzad, A., "Thickness and competition in ride-sharing markets", URL `https://ssrn.com/abstract=3065672` (2017).

Nooren, P., N. van Gorp, N. van Eijk and R. Ó. Fathaigh, "Should we regulate digital platforms? a new framework for evaluating policy options", Policy & Internet **10**, 3, 264–301 (2018).

Olivares, M. and G. P. Cachon, "Competing retailers and inventory: An empirical investigation of general motors' dealerships in isolated us markets", Management science **55**, 9, 1586–1604 (2009).

Olivares, M., C. Terwiesch and L. Cassorla, "Structural estimation of the newsvendor model: an application to reserving operating room time", Management Science **54**, 1, 41–55 (2008).

Park, J.-K. and Y. K. Ro, "The impact of a firm's make, pseudo-make, or buy strategy on product performance", Journal of Operations Management **29**, 4, 289–304 (2011).

Parker, G. and M. Van Alstyne, "Innovation, openness, and platform control", Management Science **64**, 7, 3015–3032 (2018).

Parker, G. G. and E. G. Anderson, "From buyer to integrator: The transformation of the supply-chain manager in the vertically disintegrating firm", Production and operations management **11**, 1, 75–91 (2002).

Parker, G. G. and M. W. Van Alstyne, "Two-sided network effects: A theory of information product design", Management science **51**, 10, 1494–1504 (2005).

Perols, J., C. Zimmermann and S. Kortmann, "On the relationship between supplier integration and time-to-market", Journal of Operations Management **31**, 3, 153–167 (2013).

Petrin, A. and K. Train, "A control function approach to endogeneity in consumer choice models", Journal of marketing research **47**, 1, 3–13 (2010).

Qi, W., L. Li, S. Liu and Z.-J. M. Shen, "Shared mobility for last-mile delivery: Design, operational prescriptions, and environmental impact", Manufacturing & Service Operations Management **20**, 4, 737–751 (2018).

Qin, X., Z. Liu and L. Tian, "The strategic analysis of logistics service sharing in an e-commerce platform", Omega **92**, 102153 (2020).

Qin, X., Z. Liu and L. Tian, “The optimal combination between selling mode and logistics service strategy in an e-commerce market”, European Journal of Operational Research **289**, 2, 639–651 (2021).

Rabinovich, E., “Linking e-service quality and markups: The role of imperfect information in the supply chain”, Journal of Operations Management **25**, 1, 14–41 (2007).

Rabinovich, E. and J. P. Bailey, “Physical distribution service quality in Internet retailing: Service pricing, transaction attributes, and firm attributes”, Journal of Operations Management **21**, 6, 651–672 (2004).

Rabinovich, E., A. M. Knemeyer and C. M. Mayer, “Why do Internet commerce firms incorporate logistics service providers in their distribution channels? The role of transaction costs and network strength”, Journal of Operations Management **25**, 3, 661–681 (2007).

Rahman, K. S., “From railroad to uber: Curbing the new corporate power”, URL `http://bostonreview.net/forum/k-sabeel-rahman-curbing-new-corporate-power`, accessed on April 26, 2021. (2015).

Ransom, T., “Labor market frictions and moving costs of the employed and unemployed”, Journal of Human Resources pp. 0219–10013R2 (2021).

Rey, P. and J. Tirole, “A primer on foreclosure”, Handbook of Industrial Organization **3**, 2145–2220 (2007).

Richards, T. J. and S. F. Hamilton, “Rivalry in price and variety among supermarket retailers”, American Journal of Agricultural Economics **88**, 3, 710–726 (2006).

Richards, T. J. and S. F. Hamilton, “Variety pass-through: An examination of the ready-to-eat breakfast cereal market”, Review of Economics and Statistics **97**, 1, 166–180 (2015).

Richards, T. J. and S. F. Hamilton, “Food waste in the sharing economy”, Food Policy **75**, 109–123 (2018).

Richards, T. J., S. F. Hamilton, M. Gomez and E. Rabinovich, “Retail intermediation and local foods”, American Journal of Agricultural Economics **99**, 3, 637–659 (2017).

Richards, T. J., S. F. Hamilton, K. Yonezawa and Others, “Retail market power in a shopping basket model of supermarket competition”, Tech. rep. (2015).

Rochet, J.-C. and J. Tirole, “Platform competition in two-sided markets”, Journal of the european economic association **1**, 4, 990–1029 (2003).

Rochet, J.-C. and J. Tirole, “Two-sided markets: a progress report”, The RAND journal of economics **37**, 3, 645–667 (2006).

Rosenbaum, P. R. and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects", Biometrika **70**, 1, 41–55 (1983).

Rougès, J.-F. and B. Montreuil, "Crowdsourcing delivery: New interconnected business models to reinvent delivery", in "1st International Physical Internet Conference", pp. 1–19 (2014).

Rysman, M., "Competition between networks: A study of the market for yellow pages", The Review of Economic Studies **71**, 2, 483–512 (2004).

Rysman, M., "The economics of two-sided markets", Journal of Economic Perspectives **23**, 3, 125–143 (2009).

Shapiro, C., "Protecting competition in the american economy: Merger control, tech titans, labor markets", Journal of Economic Perspectives **33**, 3, 69–93 (2019).

Shaw, J. D., J. E. Delery, G. D. Jenkins Jr and N. Gupta, "An organization-level analysis of voluntary and involuntary turnover", Academy of management journal **41**, 5, 511–525 (1998).

Singh, A., J. Zhang and S. Veeraraghavan, "Fulfillment by platform: Antitrust and upstream market power", Working Paper (2021).

Son, Y., H. E. Kwon, G. K. Tayi and W. Oh, "Impact of customers' digital banking adoption on hidden defection: A combined analytical–empirical approach", Journal of Operations Management **66**, 4, 418–440 (2019).

Soper, S., "Amazon is accused of forcing up prices in antitrust complaint", URL `https://www.bloomberg.com/news/articles/2019-11-08/amazon-merchant-lays-out-antitrust-case-in-letter-to-congress`, accessed on April 26, 2021. (2019).

Sriram, S. and M. U. Kalwani, "Optimal advertising and promotion budgets in dynamic markets with brand equity as a mediating variable", Management Science **53**, 1, 46–60 (2007).

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

Staiger, D. and J. H. Stock, "Instrumental variables regression with weak instruments", Econometrica **65**, 3, 557–586 (1997).

Steel, R. P., "Turnover theory at the empirical interface: Problems of fit and function", Academy of Management Review **27**, 3, 346–360 (2002).

Steven, A. B., Y. Dong and T. Corsi, "Global sourcing and quality recalls: An empirical study of outsourcing-supplier concentration-product recalls linkages", Journal of Operations Management **32**, 5, 241–253 (2014).

Stewart, H., D. Dong and Others, "How strong is the demand for food through direct-to-consumer outlets?", Food Policy **79**, C, 35–43 (2018).

Stock, J. H., M. Yogo *et al.*, “Testing for weak instruments in linear iv regression”, Identification and inference for econometric models: Essays in honor of Thomas Rothenberg **80**, 4.2, 1 (2005).

Straight, B., “Drivers continue switching jobs as turnover rate hits 5-year high”, URL `https://www.freightwaves.com/news/driver-issues/truck-driver-turnover-rate-increases`, accessed on August 24, 2020. (2018).

Sun, L., G. Lyu, Y. Yu and C.-P. Teo, “Fulfillment by amazon versus fulfillment by seller: An interpretable risk-adjusted fulfillment model”, Naval Research Logistics (NRL) **67**, 8, 627–645 (2020).

Swanson, A. F., “Small farmers aren’t cashing in with Wal-Mart”, URL `https://www.npr.org/sections/thesalt/2013/02/04/171051906/can-small-farms-benefit-from-wal-mart-s-push-into-local-foods`, accessed on February 13, 2020. (2013).

Ta, H., T. L. Esper and A. R. Hofer, “Designing crowdsourced delivery systems: The effect of driver disclosure and ethnic similarity”, Journal of Operations Management **60**, 19–33 (2018).

Target Corp., “Target to acquire same-Day delivery platform Shipt, Inc. to bolster fulfillment capabilities [Press Release] ”, URL `https://corporate.target.com/press/releases/2017/12/target-to-acquire-same-day-delivery-platform-shipt`, accessed on July 30, 2020. (2017).

Thilmany, D., C. A. Bond and J. K. Bond, “Going local: Exploring consumer behavior and motivations for direct food purchases”, American Journal of Agricultural Economics **90**, 5, 1303–1309 (2008).

Tian, L. and B. Jiang, “Effects of consumer-to-consumer product sharing on distribution channel”, Production and Operations Management **27**, 2, 350–367 (2018).

Tippins, M. J., K. M. Rassuli and S. C. Hollander, “An assessment of direct farm-to-table food marketing in the USA”, International Journal of Retail & Distribution Management **30**, 7, 343–353 (2002).

Toler, S., B. C. Briggeman, J. L. Lusk and D. C. Adams, “Fairness, farmers markets, and local production”, American Journal of Agricultural Economics **91**, 5, 1272–1278 (2009).

Tongarlak, M. H., D. Lee and B. Ata, “Mechanisms for increasing sourcing from capacity-constrained local suppliers”, Decision Sciences **48**, 1, 108–149 (2017).

Train, K. E., *Discrete choice methods with simulation* (Cambridge university press, 2009).

Tucker, C. and J. Zhang, “Growing two-sided networks by advertising the user base: A field experiment”, Marketing Science **29**, 5, 805–814 (2010).

US Census Bureau American Community Survey, “2017 American Community Survey”, URL https://www.census.gov/programs-surveys/acs, accessed on April 26, 2021. (2017).

US Census Bureau County Business Patterns, “2017 county business patterns”, URL https://www.census.gov/programs-surveys/cbp.html, accessed on July 30, 2020. (2016).

USDA Local Food Directories, URL https://www.ams.usda.gov/local-food-directories/foodhubs, accessed on November 26, 2020. (2020).

USDA Local Food Promotion Program, URL https://www.ams.usda.gov/services/grants/lfpp, accessed on May 20, 2021. (2021).

Vareias, A. D., P. P. Repoussis and C. D. Tarantilis, “Assessing customer service reliability in route planning with self-imposed time windows and stochastic travel times”, Transportation Science **53**, 1, 256–281 (2019).

Villas-Boas, J. M. and R. S. Winer, “Endogeneity in brand choice models”, Management science **45**, 10, 1324–1338 (1999).

Villas-Boas, J. M. and Y. Zhao, “Retailer, manufacturers, and individual consumers: Modeling the supply side in the ketchup marketplace”, Journal of Marketing Research **42**, 1, 83–95 (2005).

Villas-Boas, S. B., “Vertical relationships between manufacturers and retailers: Inference with limited data”, The Review of Economic Studies **74**, 2, 625–652 (2007).

Voccia, S. A., A. M. Campbell and B. W. Thomas, “The same-day delivery problem for online purchases”, Transportation Science **53**, 1, 167–184 (2019).

Voight, J., “As americans rush to fresh food, supermarket chains follow”, URL https://www.cnbc.com/id/49101716, accessed on February 13, 2020. (2013).

Wen, W. and F. Zhu, “Threat of platform-owner entry and complementor responses: Evidence from the mobile app market”, Strategic Management Journal **40**, 9, 1336–1367 (2019).

Weyl, E. G., “A price theory of multi-sided platforms”, American Economic Review **100**, 4, 1642–72 (2010).

Willis, D. B., C. E. Carpio and K. A. Boys, “Supporting local food system development through food price premium donations: a policy proposal”, Journal of Agricultural and Applied Economics **48**, 2, 192–217 (2016).

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

Yang, X., A. K. Strauss, C. S. Currie and R. Eglese, “Choice-based demand management and vehicle routing in e-fulfillment”, Transportation Science **50**, 2, 473–488 (2016).

Yoganarasimhan, H., “The value of reputation in an online freelance marketplace”, Marketing Science **32**, 6, 860–891 (2013).

Zervas, G., D. Proserpio and J. W. Byers, “The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry”, Journal of marketing research **54**, 5, 687–705 (2017).

Zhou, Z., L. Zhang and M. Van Alstyne, “How users create value: Platform investments that matter”, URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3625355 (2020).

Zhu, F. and Q. Liu, “Competing with complementors: An empirical look at Amazon.com”, Strategic Management Journal **39**, 10, 2618–2642 (2018).

Zurel, O. and L. Scorca, “How the fragmentation of the postal supply chain leads to new business models”, in “The Changing Postal Environment”, pp. 39–52 (Springer, 2020).

Zuurbier, P. J. P., “Supply chain management in the fresh produce industry: A mile to go?”, Journal of Food Distribution Research **30**, 856-2016-57418, 20–30 (1999).

APPENDIX A

INSTRUMENTAL VARIABLE IDENTIFICATION

We report the control function estimations for the purchase incidence and basket size models in Tables 18 and 19 as well as the first stage results for the supply provision model in Table 20.

Table 18: Control function approach first stage results-purchase incidence

	Dep.Var.: Average Price		Dep.Var.: Number of local vendors		
Variable	Estimate	Std. Err.	*Variable*	Estimate	Std. Err.
Constant	.290***	.007	Constant	.813***	.117
Wholesale Cost	.989***	.002	**Ship Volume**	-0.016***	.0003
Vacation Week *(BREAK)*	.032***	.001	**Squared Ship Volume**	.14E-04***	.65E-07
Inter-Purchase Time *(IPT)*	-0.0001***	.14E-05	Inter-Purchase Time *(IPT)*	.01***	.002
Delivery Payment Plan *(DPP)*	-0.003	.002	Delivery Payment Plan *(DPP)*	.345***	.138
Consumption Rate *(CR)*	-0.0001*	.88E-05	Consumption Rate *(CR)*	-0.066***	.003
Lagged Quantity *(LQ)*	.22E-05	.43E-05	Lagged Quantity *(LQ)*	.0002	.002
Week	.002***	.29E-05	Week	.517***	.001
Squared Week	-0.15E-05***	.72E-07	Vacation Week *(BREAK)*	-0.642***	.068
N	34,327			34,327	
F-Statistics	29887.35			31558.49	
R^2	0.87			0.88	

Table 19: Control function approach first stage results-basket size

	Dep.Var.: Average price		Dep.Var.: Number of local vendors		
Variable	Estimate	Std. Err.	*Variable*	Estimate	Std. Err.
Constant	.289***	.007	Constant	.698***	.118
Wholesale Cost	.989***	.002	**Ship Volume**	-0.016***	.0003
Lagged Quantity *(LQ)*	.19E-06	.99E-05	**Squared Ship Volume**	.14E-04***	.27E-06
Inter-Purchase Time *(IPT)*	-0.0001**	.08E-05	Lagged Quantity *(LQ)*	-0.020***	.002
Delivery Payment Plan *(DPP)*	-0.004*	.002	Inter-Purchase Time *(IPT)*	.018***	.002
Counpon Amount *(COUPON)*	.0003	.0002	Delivery Payment Plan *(DPP)*	.705***	.135
Week	.002***	.25E-05	Coupon Amount *(COUPON)*	.042***	.010
Squared Week	-0.17E-05***	.67E-07	Week	.513***	.001
Vacation Week *(BREAK)*	.032***	.001	Vacation Week *(BREAK)*	-0.624***	.069
N	34,327			34,327	
F-Statistics	29885.94			31186.52	
R^2	0.87			0.88	

Table 20: GMM first stage results

	Dep.Var.:Retail Margin			Dep.Var.: Marginal Value Local Vendors	
Variable	Estimate	Std. Err.	*Variable*	Estimate	Std. Err.
Wholesale Cost	-0.218**	.093	**One Period Lagged Marginal Value Local Vendors**	.544***	.094
Cost of Utility	.371	.359	**Two Period Lagged Marginal Value Local Vendors**	.430***	.106
Hourly Retail Wage	.210	.194	Constant	.750	.991
One Period Lagged Retail Margin	.568***	.080			
Two Period Lagged Retail Margin	.574***	.095			
Constant	-2.876	.749			
N	126			126	
F Statistics	183.64			214.59	
R^2	0.88			0.77	

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10%.

Instruments in each equation are highlighted in bold type.

APPENDIX B

ADDITIONAL PRELIMINARY EVIDENCE

Part 1 provides evidence that the routes' assignments are exogenous with respect to drivers' experience and that they do not depend on changes in base and supplementary compensations received previously by the drivers. Part 2 provides an overview of the distribution of the months of the year when drivers joined the platform as well as evidence that tenure length is not shorter among drivers who joined TForce during its annual peak demand season (from November to January). Furthermore, Part 2 provides evidence that drivers' decisions to leave the platform were not likely to occur simultaneously with decisions to leave by other drivers. Therefore, it is unlikely that drivers' decisions to leave affected other drivers' decisions to continue at the platform.

1 ROUTE ASSIGNMENTS RELATIVE TO DRIVERS' EXPERIENCE AND CHANGES IN SUPPLEMENTARY AND BASE COMPENSATIONS

According to the data, 75% of all drivers in the sample received route assignments that had no variations during their tenure at the platform. For the other 25% of drivers, weekly route assignments varied sporadically to include new neighborhoods relative to those in these drivers' assignments during the preceding week. As shown in Figure 15, route assignment variations for the drivers in the latter group took place most commonly during the drivers' early stages of tenure at the platform. For instance, while the routes originally assigned to 32 drivers upon joining TForce changed in their second week at the platform, such assignment changes applied to only 25 drivers in week 3. This number continued decreasing in drivers' tenure. By week 50, only two drivers' route assignments registered changes with respect to their assignments in the preceding week. Therefore, assignments appear to be exogenous with respect to drivers' experience at the platform since changes in these assignments become less likely as drivers accumulate more weeks at the platform.

We also examine whether route assignments for a particular week vary depending on changes in the base and supplementary payments drivers received during the preceding week. To that end, we first calculate the percentage changes in drivers' weekly base and supplementary payments relative to the payments they received in the previous week and lag these percentage changes by one period relative to each weekly route assignment. We then use a dummy variable to specify whether weekly route assignments change relative to those in the preceding week. This dummy variable takes a value of 1 if there was a change and a value of 0 otherwise. Finally, we use a logistic model to regress the dummy variable on the lagged percentage changes in the two forms of payment received by drivers. We report the results in Table 21, where Column (1) controls for month and metro area fixed effects, while Column (2) controls for month and driver fixed effects. [35] As Table 21 shows, the effects for both lagged percentage changes in base and supplementary pay are statistically non-significant at a 10% level. Therefore, we find no evidence to suggest that drivers are more likely to receive different route assignments depending on variations in weekly base and supplementary payments received previously.

[35]Note that after controlling for driver fixed effects, we must exclude those drivers who did not experience any change in their route assignments. Because a fixed-effects analysis relies only on within-driver variation, there is no within-driver variation on the dependent variable for these drivers (i.e., the dependent variable is zero every week).

Figure 15: Number of changes in route assignment by drivers' tenure

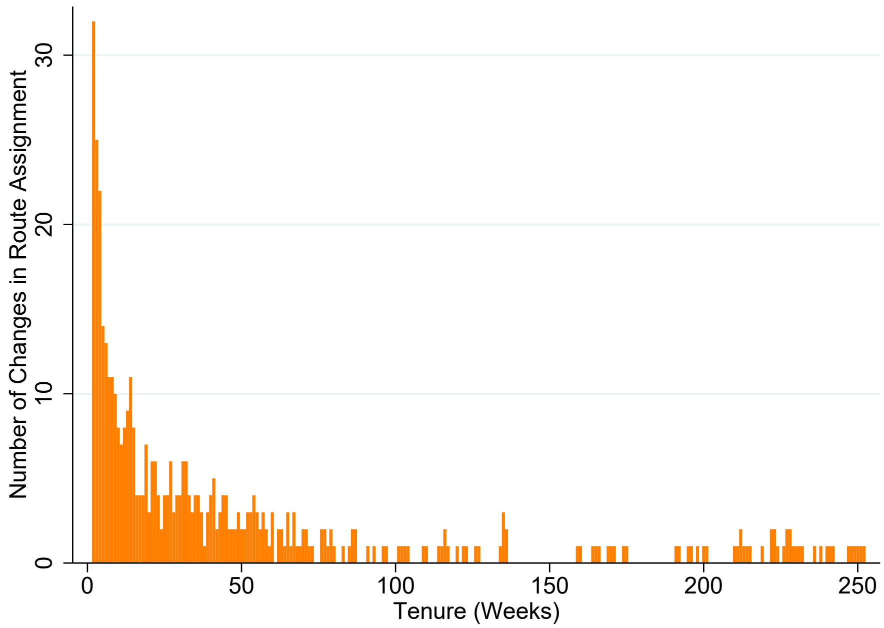


Table 21: Impact of change in compensation on route assignment

	(1)		(2)	
Dep.Var.: Change in Route Assignment	Estimate	(Std. Err.)	Estimate	(Std. Err.)
Lagged percentage change in base pay	0.000	(0.000)	0.000	(0.000)
Lagged percentage change in supplementary pay	-0.000	(0.000)	0.000	(0.000)
Observations	11,987		3,995	
Month FE	Yes		Yes	
Metro Area FE	Yes		No	
Driver FE	No		Yes	

2 DRIVER ATTRITION ANALYSIS

We first provide an overview of the distribution of the time of the year when drivers joined the platform across the nine metro areas in our study. As shown in Figure 16, drivers' starting dates at the platform are distributed fairly uniformly across all months of the year. There are variations in this distribution but they appear to depend largely on the metropolitan area where drivers work.

Figure 16: Number of drivers by starting month of the year

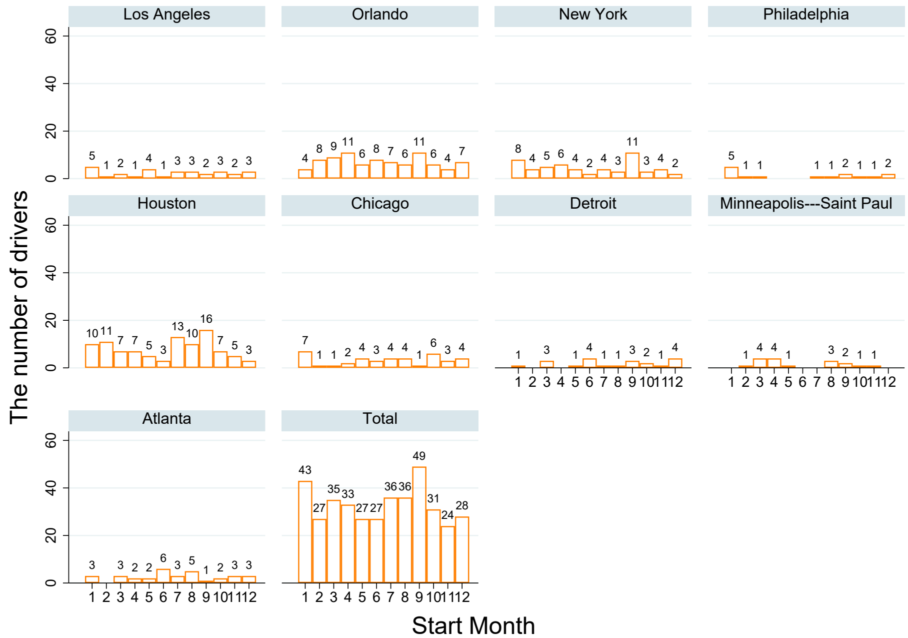
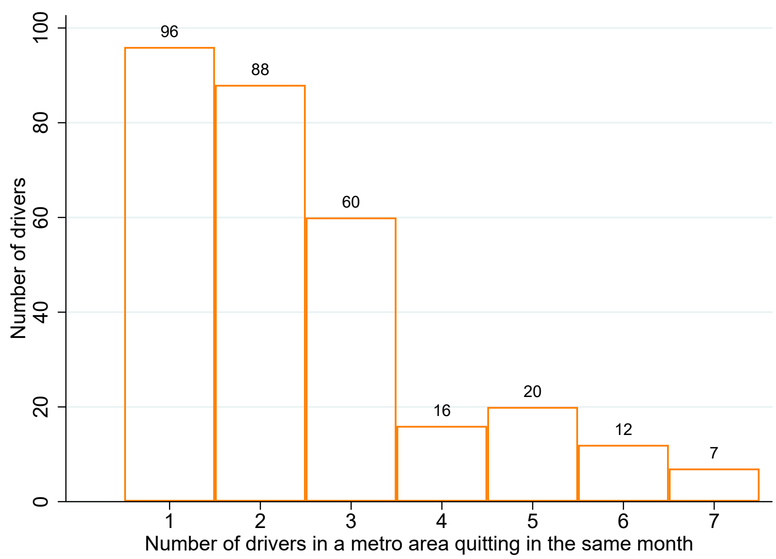


Table 22: Drivers' starting month and tenure length

Dep. Var.: Length of Tenure	Estimate	(Std Err.)
February	3.846	(10.939)
March	0.230	(10.203)
April	4.059	(10.422)
May	1.212	(10.858)
June	13.976	(11.067)
July	2.290	(9.984)
August	8.334	(10.029)
September	-6.236	(9.322)
October	26.647**	(10.393)
November	33.777***	(11.211)
December	8.697	(10.769)
Observations	396	
R-squared	0.096	
Metro Area FE	Yes	

We also investigate whether drivers who joined the platform during its annual peak demand season (from November to January) have shorter lengths of tenure. To that end, we use the length of drivers' tenure as a dependent variable and regressed it on drivers' starting months at the platform, taking January as the base. As Table 22 shows, the length of tenure for drivers who started between February and September and during the month of December is not statistically different from the length of tenure for drivers who started in January. Moreover, drivers who started in October and November have longer tenure lengths than drivers who started in January. Thus, we found no evidence to suggest that length of tenure is shorter among drivers who joined TForce during its annual peak demand season.

Figure 17: Number of drivers in a metro area leaving the platform in the same month



Finally, Figure 17 provides evidence that drivers' decisions to leave the platform were not likely to occur simultaneously with decisions to leave by other drivers. This figure includes the frequency distribution of the number of drivers in the same metro area who decided to quit the platform in the same month during our period of analysis. In total, 299 drivers quit the platform before the end of this period and, as shown in

Figure 17, 96 of these drivers quit in a month when no other drivers in the same metro area left the platform. Moreover, this figure shows that 88 drivers quit in a month when only one other driver in the same metro area also departed the platform, while 60 did so in a month when only two other drivers in the same metro area also quit. In total, 82% $\left(\frac{96+88+60}{299} \times 100\right)$ of all the drivers who left the platform during our period of analysis did so during a month in which no more than 2 other drivers in the same metro area also decided to leave the platform. Therefore, it is unlikely that drivers' decisions to leave affected other drivers' decisions to continue at the platform. This is not very surprising, particularly because any influence on these decisions is limited to drivers working in the same metro area and, in each of these areas, drivers work independently serving routes typically assigned to them when they join the platform.

APPENDIX C

DERIVATION OF CONDITIONAL CHOICE PROBABILITIES ESTIMATOR

Recall that we defined the conditional value function as $v_k(X_{ijt}) \equiv V_k(X_{ijt}) - \varepsilon_{ijkt}$. Then, we can rewrite Equation (18) in Section 4 as

$$\begin{aligned} V_1(X_{ijt}) &= u_1(X_{ijt}) + \varepsilon_{ij1t} \\ &+ \beta E\left[p_0(X_{ijt+1})(v_0(X_{ijt+1}) + \tilde{\varepsilon}_{ij0t+1}) + p_1(X_{ijt+1})(v_1(X_{ijt+1}) + \tilde{\varepsilon}_{ij1t+1}))\right] \\ &= u_1(X_{ijt}) + \varepsilon_{ij1t} + \beta E[p_0(X_{ijt+1})\tilde{\varepsilon}_{ij0t+1} + p_1(X_{ijt+1})\tilde{\varepsilon}_{ij1t+1} \\ &+ p_1(X_{ijt+1})(v_1(X_{ijt+1}) - v_0(X_{ijt+1})) + v_0(X_{ijt+1})], \end{aligned} \tag{35}$$

where the second equality exploits the relationship that $p_0(X_{ijt+1}) = 1 - p_1(X_{ijt+1})$. Using the relationship shown in Equation (22) in Section 4 that the difference between value functions $v_1 - v_0$ is equal to $\log\left(\dfrac{p_1}{1-p_1}\right)$ and applying the law of iterated expectations, Equation (35) becomes

$$\begin{aligned} V_1(X_{ijt}) &= u_1(X_{ijt}) + \varepsilon_{ij1t} + \beta E_{X_{ijt+1}}\Big[p_0(X_{ijt+1})\tilde{\varepsilon}_{ij0t+1} + p_1(X_{ijt+1})\tilde{\varepsilon}_{ij1t+1} \\ &+ p_1(X_{ijt+1})\log\left(\frac{p_1(X_{ijt+1})}{1-p_1(X_{ijt+1})}\right) + v_0(X_{ijt+1})|X_{ijt+1}\Big] \\ &= u_1(X_{ijt}) + \varepsilon_{ij1t} + E_{X_{ijt+1}}\Big[V_0(X_{ijt}) + p_0(X_{ijt+1})E[\tilde{\varepsilon}_{ij0t+1}|X_{ijt+1}, d^*_{t+1} = 0] \\ &+ p_1(X_{ijt+1})E[\tilde{\varepsilon}_{ij1t+1}|X_{ijt+1}, d^*_{t+1} = 1] + p_1(X_{ijt+1})\log\left(\frac{p_1(X_{ijt+1})}{1-p_1(X_{ijt+1})}\right)\Big]. \end{aligned} \tag{36}$$

Next, we derive the conditional expectation of Gumbel random variables $E\left[\tilde{\varepsilon}_{ij1t+1}|X_{ijt+1}, d^*_{t+1}\right]$ and $E\left[\tilde{\varepsilon}_{ij0t+1}|X_{ijt+1}, d^*_{t+1}\right]$. Recall that ε_{ijkt} are i.i.d. Gumble variables with location and scale parameters 0 and 1. The probability density and cumulative distribution functions of ε_{ijkt} are

$$g(z) = e^{-z}e^{-e^{-z}}, z \in (-\infty, \infty)$$

$$G(z) = e^{-e^{-z}}, z \in (-\infty, \infty).$$

Therefore,

$$\begin{aligned}
E\left[\tilde{\varepsilon}_{ij1t+1}|X_{ijt+1},d^*_{t+1}\right] &= \frac{1}{p_1}\int_{-\infty}^{\infty}\varepsilon_1 g(\varepsilon_1)G(\varepsilon_1+v_1-v_0)d\varepsilon_1 \\
&= \frac{1}{p_1}\int_{-\infty}^{\infty}\varepsilon_1 g(\varepsilon_1)G\left(\varepsilon_1+\log\left(\frac{p_1}{1-p_1}\right)\right)d\varepsilon_1 \\
&= \frac{1}{p_1}\int_{-\infty}^{\infty}\varepsilon_1 e^{-\varepsilon_1}e^{-e^{-\varepsilon_1}}e^{-e^{-\left(\varepsilon_1+\log\left(\frac{p_1}{1-p_1}\right)\right)}}d\varepsilon_1 \\
&= \frac{1}{p_1}\int_{-\infty}^{\infty}\varepsilon_1 e^{-\varepsilon_1}e^{-e^{-\varepsilon_1\left(1+e^{-\left(\log\left(\frac{p_1}{1-p_1}\right)\right)}\right)}}d\varepsilon_1 \\
&= \frac{1}{p_1}\int_{-\infty}^{\infty}\varepsilon_1 e^{-\varepsilon_1}e^{-e^{-\varepsilon_1\left(1+\left(\frac{p_1}{1-p_1}\right)\right)}}d\varepsilon_1 \\
&= \int_{-\infty}^{\infty}\varepsilon_1\frac{e^{-\varepsilon_1}}{p_1}e^{\frac{e^{-\varepsilon_1}}{p_1}}d\varepsilon_1.
\end{aligned}$$

Note the second equality exploits the fact that v_1-v_0 is equal to $\log\left(\frac{p_1}{1-p_1}\right)$. Next, we let $x=\frac{e^{-\varepsilon_1}}{p_1}$. Then $\varepsilon_1=-\log(p_1x)$, $dx=-\frac{e^{-\varepsilon_1}}{p_1}d\varepsilon_1$, $\varepsilon_1=-\infty \implies x=\infty$, $\varepsilon_1=\infty \implies x=0$. Therefore, by change of variables

$$E\left[\tilde{\varepsilon}_{ij1t+1}|X_{ijt+1},d^*_{t+1}\right]=\int_0^{\infty}\log(p_1x)e^{-x}dx=-\int_0^{\infty}(\log(p_1)+\log(x))e^{-x}dx=\gamma-\log p_1, \tag{37}$$

where γ is Euler constant (≈ 0.577), the mean of a standard Type I extreme value distribution (McFadden, 1974).

Similarly,

$$\begin{aligned}
E\left[\tilde{\varepsilon}_{ij0t+1}|X_{ijt+1},d^*_{t+1}\right] &= \frac{1}{p_0}\int_{-\infty}^{\infty}\varepsilon_1 g(\varepsilon_0)G(\varepsilon_0+v_0-v_1)d\varepsilon_0 \\
&= \frac{1}{p_1}\int_{-\infty}^{\infty}\varepsilon_0 g(\varepsilon_0)G\left(\varepsilon_0+\log\left(\frac{p_0}{1-p_0}\right)\right)d\varepsilon_0 \qquad (38) \\
&= \gamma-\log p_0.
\end{aligned}$$

Note that, we can write $V_0(X_{ijt})$ introduced in Equation (19) in Section 4 as

$$\begin{aligned}
V_0(X_{ijt}) &= u_0(X_{ijt}) + \varepsilon_{ij0t} + \sum_{s=t+1}^{T} \beta^{s-t} E\left[u_0(X_{ijs}) + \tilde{\varepsilon}_{ij0s}\right] \\
&= \varepsilon_{ij0t} + E[\tilde{\varepsilon}_{ij0s}] \sum_{s=t+1}^{T} \beta^{s-t} \\
&= \gamma \left(\frac{\beta - \beta^{T-t}}{1-\beta} \right),
\end{aligned} \tag{39}$$

then, $V_0(X_{ijt+1})$ as

$$V_0(X_{ijt+1}) = \gamma \left(\frac{\beta - \beta^{T-t-1}}{1-\beta} \right). \tag{40}$$

Next, we substitute (37) - (40) into (36) and integrate over possible states in week $t+1$

$$\begin{aligned}
V_1(X_{ijt}) &= u_1(X_{ijt}) + \varepsilon_{ij1t} \\
&+ \beta E_{X_{ijt+1}} \Bigg[\gamma \left(\frac{\beta - \beta^{T-t-1}}{1-\beta} \right) + (1 - p_1(X_{ijt+1}))[\gamma - \log(1 - p_1(x_{ijt+1}))] \\
&+ p_1(X_{ijt+1}) \left[\gamma - \log(p_1(X_{ijt+1})) + \log\left(\frac{p_1(X_{ijt+1})}{1 - p_1(X_{ijt+1})} \right) \right] \Bigg] \\
&= u_1(X_{ijt}) + \varepsilon_{ij1t} + \beta E_{X_{ijt+1}} \left[\gamma \left(\frac{\beta - \beta^{T-t-1}}{1-\beta} \right) + \gamma - \log(1 - p_1(X_{ijt+1})) \right] \\
&= u_1(X_{ijt}) + \varepsilon_{ij1t} + \beta \left(\gamma \left(\frac{1 - \beta^{T-t-1}}{1-\beta} \right) - \log(p_0(X_{ijt})) f(X_{ijt+1}|X_{ijt}) dX_{ijt+1} \right).
\end{aligned} \tag{41}$$

Finally, we take the difference between choice-specific conditional value function using Equation (41) and Equation (39) to obtain Equation (22) as shown in Section 4

$$\begin{aligned}
v_1(X_{ijt}) - v_0(X_{ijt}) &= V_1(X_{ijt}) - V_0(X_{ijt}) - \varepsilon_{ij1t} + \varepsilon_{ij0t} \\
&= u_1(X_{ijt}) - \beta \int \log(p_0(X_{ijt})) f(X_{ijt+1}|X_{ijt}) dX_{ijt+1}.
\end{aligned} \tag{42}$$

APPENDIX D

DETAILS OF FIRST-STEP ESTIMATION

We assume the state transition probabilities follow an AR(1) process. The forecasting shocks to the one-period ahead state variables, denoted by $\varsigma_t^{(\cdot)}$, are assumed to be normally distributed and independent over time. The state variables of efforts include miles per stop and number of hours. When drivers forecast their future expected density of routes, D_{ijt}, they use the last week's density of routes, $D_{ij,t-1}$, tenure T_{ijt}, metro area fixed effects ζ_j and month fixed effects ϕ_t. A similar specification is used for the forecast of hours, H_{ijt}.

$$D_{ijt} = \sigma_{D1} + \sigma_{D2} D_{ij,t-1} + \sigma_{D3} T_{ijt} + \zeta_j + \phi_t + \varsigma_t^D \tag{43}$$

$$H_{ijt} = \sigma_{H1} + \sigma_{H2} H_{ij,t-1} + \sigma_{H3} T_{ijt} + \zeta_j + \phi_t + \varsigma_t^H. \tag{44}$$

We assume that drivers forecast their base pay $\log W_{ijt}$ based on the last week's $\log W_{ij,t-1}$, tenure T_{ijt}, squared tenure T_{ijt}^2, metro area fixed effects ζ_j and month fixed effects ϕ_t. Following a standard Mincerian wage equation (Mincer, 1974), we use logged wage and introduce the squared term of tenure to capture the concave shape of experience-productivity profile. The transition of supplementary pay ($\log I_{ijt}$) has a similar specification except for not having the squared term. The reason is that supplementary pay is expected to decrease in tenure monotonically.

$$\log W_{ijt} = \sigma_{W1} + \sigma_{W2} \log W_{ij,t-1} + \sigma_{W3} T_{ijt} + \sigma_{W4} T_{ijt}^2 + \zeta_j + \phi_t + \varsigma_t^W \tag{45}$$

$$\log I_{ijt} = \sigma_{I1} + \sigma_{I2} \log I_{ij,t-1} + \sigma_{I3} T_{ijt} + \zeta_j + \phi_t + \varsigma_t^I. \tag{46}$$

Table 23 shows the empirical results of the state transition probability functions using the specifications in Equations (43) to (46). Table 24 reports results from a flexible logit model to estimate the conditional choice probabilities of quitting in the first stage (see Section 5 .1).

Table 23: Transition probability

Parameter	Description	Estimate	(Std. Err.)
	Log base pay		
Lag log base pay	One week lagged of log base pay	0.635***	(0.006)
Tenure (week/10)		0.018***	(0.005)
Tenure (week/10) ^{2}		-0.0007**	(0.0003)
Constant		2.364***	(0.056)
R^2			0.511
	Log supplement pay		
Lag log supplement pay	One week lagged of log supplement pay	0.725***	(0.006)
Tenure (week/10)		-0.013***	(0.002)
Constant		1.215***	(0.057)
R^2			0.434
	Hours		
Lag hours (/10)	One week lagged of hours worked	0.593***	(0.007)
Tenure (week/10)		-0.006***	(0.002)
Constant		1.425 ***	(0.042)
R^2			0.436
	Mile per stop		
Lag mile per stop	One week lagged of mile per stop	0.701***	(0.006)
Tenure (week/10)		-0.001	(0.004)
Constant		0.775***	(0.065)
R^2			0.515

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) metro area and month fixed effects are included for all transition equations and results are omitted here; (4) hours and tenure are scaled down by a factor of 1/10.

Table 24: Conditional choice probability

Variables	Estimate	(Std. Err.)
Constant	-2.334***	(0.389)
Base pay	-0.531***	(0.101)
Base pay^2	0.013***	(0.003)
Supplementary pay	-0.338**	(0.146)
Supplementary pay^2	0.034***	(0.011)
Hours	0.519*	(0.283)
Hours^2	-0.033	(0.069)
Miles per stop	0.280***	(0.055)
Miles per stop^2	-0.007***	(0.002)
Base pay × Supplementary pay	0.033***	(0.010)
Base pay × Hours	-0.009	(0.034)
Base pay × Miles per stop	0.014**	(0.006)
Supplementary pay × Hours	-0.083*	(0.049)
Miles per stop × Hours	-0.090***	(0.025)
Miles per stop × Supplementary pay	-0.031*	(0.018)
Tenure	-0.153**	(0.062)
Tenure^2	-0.001	(0.004)
Base pay × Tenure	0.006	(0.008)
Supplementary pay × Tenure	0.029**	(0.014)
Hours× Tenure	0.003	(0.026)
Miles per stop × Tenure	-0.000	(0.005)
LL	-1195.070	

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) metro area and month fixed effects are included for all transition equations and results are omitted here; (4) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10; (5) conditional choice probabilities are based on the probability of quitting (the terminal choice)

APPENDIX E

DETAILS OF SECOND-STEP ESTIMATION

All the estimates of the state transition probability functions and conditional choice probability function are known at this point. Using these estimates, we can now simulate the value of the one-period ahead conditional choice probabilities. We then present the estimation process first without considering unobserved heterogeneity and then while taking unobserved heterogeneity into account.

Step 1 Simulate the one period ahead state variables: $X_{ijt} \in (W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt})$:

$$D_{ijt} = \sigma_{D1} + \sigma_{D2} D_{ij,t-1} + \sigma_{D3} T_{ijt} + \zeta_j + \phi_t + \varsigma_t^D$$

$$H_{ijt} = \sigma_{H1} + \sigma_{H2} H_{ij,t-1} + \sigma_{H3} T_{ijt} + \zeta_j + \phi_t + \varsigma_t^H$$

$$W_{ijt} = (W_{ij,t-1})^{\sigma_{W2}} \, e^{\sigma_{W1} + \sigma_{W3} T_{ijt} + \sigma_{W4} (T_{ijt})^2 + \zeta_j + \phi_t + \varsigma_t^W}$$

$$I_{ijt} = (I_{ij,t-1})^{\sigma_{I2}} \, e^{\sigma_{I1} + \sigma_{I3} T_{ijt} + \zeta_j + \phi_t + \varsigma_t^I},$$

where $(\tilde{\varsigma}_t^D, \tilde{\varsigma}_t^H, \tilde{\varsigma}_t^W, \tilde{\varsigma}_t^I)$ are the iid standard normal random variables reflecting the empirical distribution of their corresponding state variables.

Step 2 without unobserved heterogeneity:

Calculate $\int \log \left[p_0 \left(X_{ijt+1} \right) \right] f \left(X_{ijt+1} | X_{ijt} \right) dX_{ijt+1}$ by simulation similar to the approach used in Huang and Smith (2014), Murphy (2018), and Ransom (2021). We integrate the logged one-period ahead conditional choice probabilities over the empirical distribution of state variables using Monte Carlo methods:

$$\int \log p_0 \left(\varsigma_{t+1}^D, \varsigma_{t+1}^H, \varsigma_{t+1}^W, \varsigma_{t+1}^I \right) dF \left(\varsigma^D, \varsigma^H, \varsigma^W, \varsigma^I \right) \approx \frac{1}{D} \sum_{d=1}^{D} \log p_0 \left(\varsigma_d^D, \varsigma_d^H, \varsigma_d^W, \varsigma_d^I \right), \tag{47}$$

where $\left(\varsigma_{t+1}^D, \varsigma_{t+1}^H, \varsigma_{t+1}^W, \varsigma_{t+1}^I \right)$ are the vectors of shocks to the forecasting of their corresponding state in week $t+1$. We draw D draws from the standard normal distribution, plugging them into the conditional choice probabilities. We then take the average of the conditional choice probabilities value and multiply it by β. Next, we can take $p_0\left(X_{ijt+1}\right)$ as given and estimate the structural parameters using a simple logistic regression specification.

We now turn into the estimation that considers unobserved heterogeneity. In this case, the step 1 remains the same, but step 2 is more involved which is discussed as follows:

Step 2 with unobserved heterogeneity: Recall that type s_i be a dummy variable which takes 0 ($s_i = r - 1 = 0$) when a driver belongs to the first category, and 1 ($s_i = r - 1 = 1$) when she belongs to the second category. Then, we let the initial guess of population probabilities be $\pi^1 = \{\pi_1^1, \pi_2^1\} = \{0.5, 0.5\}$ and the starting values of structural parameters $\theta = (\theta_1, \ldots, \theta_5)$ be the estimates obtained from the estimation without controlling for unobserved heterogeneity.

Step 3 Update the probability of a driver being a certain type q_{ir}:

$$q_{ir}^2 = \frac{\pi_r^1 \prod_{t=1}^T \mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta^1, \pi^1, s_i = r-1)}{\sum_{r=1}^2 \pi_r^1 \prod_{t=1}^T \mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta^1, \pi^1, s_i = r-1)} \ \forall\ r. \tag{48}$$

Then, we average q_{ir} over the total number of drivers to update the population probabilities of types:

$$\pi_r^2 = \frac{\sum_{i=1}^I q_{ir}^2(r|X_{ijt}, \theta^1, \pi^1)}{I} \ \forall\ r. \tag{49}$$

Step 4: Numerically update $\beta \int \log\left[p_0\left(X_{ijt+1}, s_i\right)\right] f\left(X_{ijt+1}|X_{ijt}\right) dX_{ijt+1}$ for each type r similar to the approach used in *Step 2 without unobserved heterogeneity.* Next, plug this value into the choice probability and update the structural parameters by maximizing:

$$\theta^2 = \arg\max \sum_{i=1}^I \sum_{t=1}^T \sum_{r=1}^2 q_{ir}^2 \times log\left[\mathcal{L}_{it}(d_{it} \mid X_{ijt}, \theta^1, s_i = r-1)\right]. \tag{50}$$

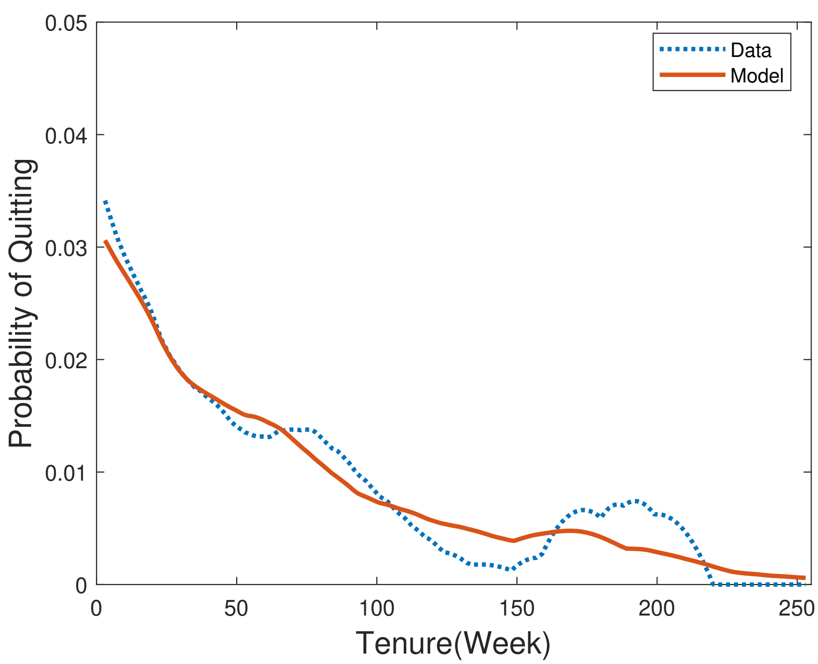
Finally, use the new estimates of θ^2, go back to step 2 with unobserved heterogeneity and update the q_{ir} and continue to repeat steps 3-4 until the estimates of θ and π converge.

APPENDIX F

ROBUSTNESS CHECKS AND MODEL FIT

Below, we evaluate the dynamic choice model's two-stage estimation performance by comparing the predicted value of quitting with the realized value of quitting observed in the data. To that end, we used the estimates of the structural parameters in Column (1) in Table 9 to simulate 1,000 times the predicted value of quitting. We then plotted the predicted quit hazard with the data-based quit hazard using an Epanechnikov kernel with a bandwidth of 20 weeks (see Figure 18). Both the model-predicted and data-based quit hazard are decreasing in tenure. Although the hazard rates as a function of tenure are consistent across, the model predicts the quit hazard better in the early stages of tenure than in the latter periods. This is likely driven by the smaller number of observations in the latter periods because most drivers do not stay at the platform for more than 50 weeks (as shown in Table 6, 260 out of 396 drivers quit within the first 50 weeks of their tenure).

Figure 18: Model fit: Quit hazard



We also present the results of a test we conducted to evaluate the robustness of our main results relative to the discount factor. Therefore, we used two alternative values

Table 25: Structural estimation results with types and discount factors

	Discount factor=0.9 (1)		Discount factor=0.8 (2)	
	Estimate	(Std. Err.)	Estimate	(Std. Err.)
Type	0.332**	(0.141)	0.385***	(0.141)
Constant	0.494	(0.359)	0.658*	(0.358)
Base pay (week/$100)	0.130***	(0.029)	0.134***	(0.029)
Supplementary pay(week/$100)	0.103**	(0.047)	0.111**	(0.047)
Hours (/10)	0.010	(0.079)	0.041	(0.079)
Miles per stop	-0.074***	(0.015)	-0.073***	(0.015)
Tenure (week/10)	0.033	(0.022)	0.040*	(0.022)
Type 1	0.652		0.663	
Type 2	0.348		0.337	
LL	-1225.059		-1225.656	
obs	15293		15293	
Number of drivers	396		396	

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10.

(0.90 and 0.80) to replace the discount factor of 0.9957 we used to generate these results. As shown in Table 25, the signs and magnitudes of the parameter estimates remain largely consistent regardless of the value chosen for the discount factor.

Next, we address two potential concerns over the results of the structural estimation in Column (1) in Table 9. The first concern is that Table 9's structural estimates of state variables such as tenure and hours are not statistically significant. However, they are still included in the first stage to obtain conditional choice probabilities. To address this concern, we exclude tenure and hours from the state vector and re-estimate the model. The results from this alternative model (in Tables 26 and 27) remain consistent with the main results. However, the evaluation of the alternative model's performance reveals that the model does not predict the value of quitting well especially during the latter stages of tenure, as shown in Figure 19. This underscores the importance of controlling for hours and tenure in predicting drivers' probability of quitting.

Table 26: Structural estimation results with types excluding tenure and hours as state variables

Variables	Estimate	(Std. Err.)
Constant	0.093	(0.356)
Base pay (week/$100)	0.139***	(0.016)
Supplementary pay(week/$100)	0.104**	(0.044)
Miles per stop	-0.048***	(0.016)
Type	0.332**	(0.144)
Type 1	0.659	
Type 2	0.341	
LL	-1210.052	
obs	15293	
Number of drivers	396	

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10.

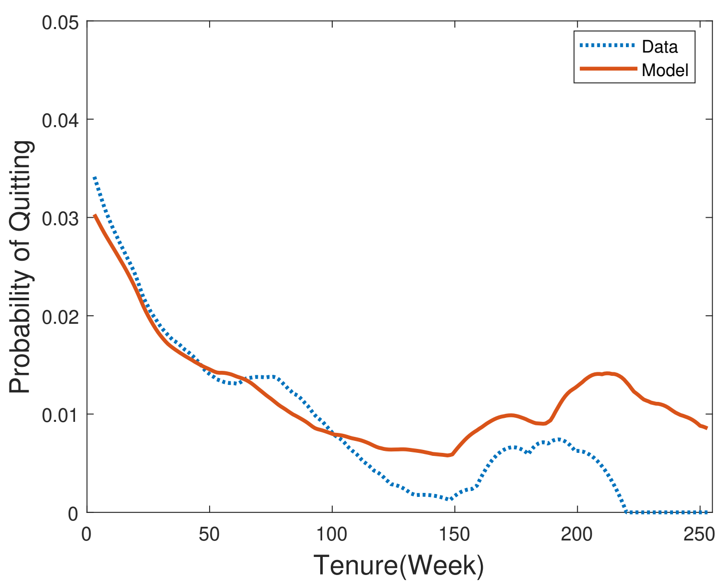
The other concern is that not all the estimates of conditional choice probabilities are statistically significant and using these estimates to simulate conditional choice probabilities may lead to overfitting. Conditional choice probabilities are inverted as a selection correction term in the second stage estimation to adjust for the fact that the choice made may not be optimal (Ellickson *et al.*, 2012; De Groote and Verboven, 2019). In the past, studies have predicted these probabilities by including all estimates, both significant and non-significant, obtained from flexible specifications (e.g., Arcidiacono and Miller, 2011; Yoganarasimhan, 2013; Chung *et al.*, 2014). Although our analysis is consistent with this approach, we evaluated the robustness of our results by excluding the non-significant parameters when simulating the conditional choice probabilities.

Table 27: Summary statistics for the two types excluding tenure and hours as state variables

	Type 1	Type 2
Type	0.659	0.341
Base pay (week)	819.235	822.415
Supplementary pay (week)	156.578	123.738
Hours	32.755	30.692
Miles per stop	3.426	3.295
Length of Tenure (week)	21.344	84.685

Notes: The population probability of each type is the estimated value for π_r; drivers are classified into a type based on the value estimated for q_{ir} (i.e., driver i is type 1 if $q_{i1} > 0.5$).

Figure 19: Model fit of quit hazard excluding tenure and hours



Therefore, we excluded all estimates of the flexible logit model that are statistically non-significant along with the non-significant estimates of metro area and month fixed effects. Table 28 presents the results. The estimates in Table 28 are consistent with those in Table 9. The only difference is that type is not statistically significant.

Table 28: Structural estimation results with types excluding non-significant estimates from simulated conditional choice probabilities

Variables	Estimate	(Std. Err.)
Constant	0.096	(0.360)
Base pay (week/$100)	0.117***	(0.028)
Supplementary pay(week/$100)	0.081*	(0.047)
Hours (/10)	-0.054	(0.077)
Miles per stop	-0.076***	(0.015)
Tenure (week/10)	0.085***	(0.021)
Type	0.209	
	(0.148)	
Type 1	0.523	
Type 2	0.477	
LL	-1239.830	
obs	15293	
Number of drivers	396	

Notes: (1) Standard errors are in parentheses; (2) *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively; (3) to avoid numerical overflows caused by large values, base pay and supplementary pay are scaled down by a factor of 1/100 and hours and tenure are scaled down by a factor of 1/10.

APPENDIX G

ENDOGENEITY ISSUES AND INSTRUMENTAL VARIABLES

In this appendix, we address potential endogeneity concerns in our setting. The endogeneity issues are primarily caused by the base payments that compensate drivers for their productivity. In particular, the base payments depend on the number of parcels delivered and drivers of high ability are more likely to complete deliveries for a high number of parcels which yields a large base payment. The endogeneity arises because drivers' ability is unobserved to us and, more importantly, it tends to be correlated with their decisions to quit or stay at the platform. For example, drivers of high ability are also those with attractive outside options and hence are more likely to quit the platform. Consequently, overlooking this issue may lead to a downward biased estimate of the effect of the base payments on driver retention. To account for the potential endogeneity of the base payments, we employ a control function approach with instrumental variables (IVs).

Ideal IVs for the base payments should be highly correlated with the base payments and affect drivers' likelihood of staying only through the base payments. Following Stafford (2015), our IV strategy is based on the exogenous weekly variations in weather in the areas where drivers performed deliveries. The idea is that the fluctuations in weather conditions have an impact on the difficulty of making deliveries which in turn affects the number of parcels delivered by drivers and hence their base payments. On the other hand, conditional on the base payments, the fluctuations in weather conditions are unlikely to affect drivers' decisions to quit or stay at the platform. To that end, we collected county-level daily weather measures from the U.S. National Climatic Data Center. [36] The key measures of weather consist of the daily maximum and minimum temperature and daily precipitation. We take the average of the daily values for each weather measure over a week period to obtain its corresponding weekly

[36] The data can be accessed at the website of the National Oceanic and Atmospheric Administration: http://www.ncdc.noaa.gov/oa/ncdc.html.

value. We then assign the values of weekly maximum and minimum temperature and weekly precipitation to the driver-week observations based on the county where the drivers made deliveries in a week.

We carefully evaluate the validity of weather as IVs to ensure they satisfy both the inclusion and exclusion restrictions. To validate the inclusion restriction, we test for the strength of the IVs using an F-test from the first-stage IV regression (Staiger and Stock, 1997; Stock *et al.*, 2005). As shown in Table 29, the F-statistic (684.24) and R^2 values (0.538) obtained in the first-stage IV regression of the base payments on maximum and minimum temperature, precipitation, and the exogenous explanatory variables in the utility of staying suggest that these IVs are not weak. Second, the exclusion restriction ensures that the IVs are not correlated with the error term in drivers' utility of staying. In our setting, weather conditions are exogenous transitory shocks that are unlikely to influence drivers' permanent decision to quit or stay at the platform. This is in line with the seminal work of Farber (2005), studying the labor supply decisions of taxi drivers, which shows that weather shocks, such as rainfall, do not directly explain drivers' probability of stopping driving in a day.

Our two-stage approach in estimating the structural parameters makes it possible to adopt an IV method to address the potential endogeneity. Although IV methods are widely used to address endogeneity in static models, it is not the case in dynamic models. The main reason is that the traditional approach to estimating dynamic models requires the computation of continuation value functions by fully solving the dynamic problem. When the continuation values depend on an endogenous state variable, the use of IV methods can further increase complexity and computational burden, making the dynamic model intractable (Kalouptsidi *et al.*, 2021). As we use the two-stage approach that obviates the need to fully solve the dynamic problem, we are able to apply the IV method in a tractable way into our dynamic model. In

Table 29: Control function approach first stage results

Dep.Var.: Base pay (week/$100)		
Variables	Estimate	(Std. Err.)
Max. temperature	0.058***	(0.007)
Min. temperature	-0.073***	(0.007)
Precipitation	0.030	(0.134)
Supplementary pay (week/$100)	-0.342***	(0.016)
Hours /10	2.338***	(0.020)
Miles per stop	0.023***	(0.008)
Tenure (week/10)	0.128***	(0.005)
obs	15,293	
R^2	0.538	
F-statistic	684.24	

Notes: (1) Instruments are highlighted in bold type. (2) Both month and metro area fixed effects are included in the first stage regression and the estimates of the fixed effects are omitted in this table.

particular, our IV approach is similar to those used in the recent studies that explore the two-stage framework (e.g., De Groote and Verboven, 2019; Diamond *et al.*, 2019). In the first stage, we simulate the state transition probability functions, the conditional choice probability function, and the probability of a driver being a type as detailed in Section 5 . In this stage, we also implement the control function method by regressing the base payments on the IVs to obtain the residuals (Imbens and Wooldridge, 2007; Wooldridge, 2015). In the second stage, we include the residuals as controls in the utility of staying to estimate the structural parameters.

We report the structural parameter estimates obtained after controlling for endogeneity of the base payments in Table 30. We find that the coefficient of base payments is of the expected sign and statistically significant ($p < 0.01$). Moreover, the coefficient is higher compared to that obtained without controlling for this endogeneity as reported in Column (1) of Table 9 (0.470 vs. 0.127). This indicates a downward bias in the estimated coefficient for the effect of base payments on drivers' likelihood of staying when no endogeneity is taken into account. The coefficient of the control function is negative and statistically significant (-0.344, $p < 0.05$), suggesting that the unobserved terms are negatively correlated with drivers' likelihood of staying. Both results point to evidence that our estimated value for the effect of base payments on drivers' retention in the main estimation represents a conservative lower bound. The values of the coefficients for the remaining state variables are highly consistent with those obtained without control function, except for the constant term which was non-significant without control function but is now negative and statistically significant (-2.122, $p < 0.01$).

Table 30: Estimation results with control function

Variables	Estimate	(Std. Err.)
Type	0.320**	(0.140)
Constant	-2.122**	(1.053)
Base pay (week/100 $)	0.470***	(0.143)
Supplementary pay (week/100 $)	0.099**	(0.047)
Hours /10	-0.009	(0.078)
Miles per stop	-0.075***	(0.015)
Tenure	0.021	(0.023)
Control function	-0.344**	(0.141)
LL	-1207.054	
obs	15293	
Number of drivers	396	

Standard errors are in parentheses, *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.

APPENDIX H

PROOFS OF PROPOSITIONS

Proof of Proposition 1. We differentiate with respect to $x \in \{W_{ijt}, I_{ijt}, H_{ijt}, D_{ijt}, T_{ijt}\}$ to obtain

$$\frac{\partial p_1\left(X_{ijt}\right)}{\partial x} = \frac{e^{-u_1\left(X_{ijt}\right)}}{\left(1+e^{-u_1\left(X_{ijt}\right)}\right)^2} \frac{\partial u_1}{\partial x} = p_0\left(X_{ijt}\right) p_1\left(X_{ijt}\right) \frac{\partial u_1}{\partial x}.$$

The first term is positive, and thus the sign of $\frac{\partial p_1\left(X_{ijt}\right)}{\partial x}$ is determined by the sign of $\frac{\partial u_1}{\partial x}$. □

Proof of Proposition 2. Let

$$q\left(X_{ijt}\right) = \frac{1}{\left(1+e^{v_0\left(X_{ijt}\right)-v_1\left(X_{ijt}\right)}\right)}.$$

Recall that

$$v_1\left(X_{ijt}\right) = u_1 + \sum_{s=t+1}^{T} \beta^{s-t} E[v_1(X_{ijt+1})], \tag{51}$$

and that $v_1^{(1)}(= v_1 - u_1)$ is defined as the conditional value function from period $t+1$ onward conditional a decision to continue in period t. Substituting it into (27) and implicitly differentiating with respect to x

$$\begin{aligned} \frac{\partial p_1\left(X_{ijt}\right)}{\partial x} = \frac{\partial q\left(X_{ijt}\right)}{\partial x} &= \frac{\partial q}{\partial(v_0-v_1)} \frac{\partial(v_0-v_1)}{\partial x} \\ &= \frac{\partial q}{\partial(v_0-v_1)} \left(\frac{\partial v_0}{\partial x} - \frac{\partial v_1^{(1)}}{\partial x} - \frac{\partial u_1}{\partial x}\right). \end{aligned} \tag{52}$$

Note that

$$\frac{\partial q}{\partial(v_0-v_1)} = -\frac{e^{v_0-v_1}}{(1+e^{v_0-v_1})^2} < 0. \tag{53}$$

Recall that v_0 is the expected optimal utility in weeks t through T given a decision to quit in the current week t without the error term, i.e., $v_0(X_{ijt}) = \gamma\left(\frac{\beta-\beta^{T-t}}{1-\beta}\right)$ (see (39)). Therefore,

$$\frac{dv_0}{dx} = 0. \tag{54}$$

Substituting into the above to obtain part (i),

$$\frac{\partial p_1\left(X_{ijt}\right)}{\partial x}=\frac{e^{v_0(X_{ijt})-v_1(X_{ijt})}}{\left(1+e^{v_0(X_{ijt})-v_1(X_{ijt})}\right)^2}\left(\frac{\partial u_1}{\partial x}+\frac{\partial v_1^{(1)}}{\partial x}\right)=p_0\left(X_{ijt}\right)p_1\left(X_{ijt}\right)\left(\frac{\partial u_1}{\partial x}+\frac{\partial v_1^{(1)}}{\partial x}\right).$$

Suppose that $x=W_{ijt}$ so that the presentation of the following arguments is less abstract (the arguments similarly apply to other values of x). Recall that, φ_W is a function of W_{ijt}, tenure T_{ijt}, and the iid standard normal random variables ς^W capturing forecasting shocks to the one-period ahead state variable W_{ijt+1} (see Appendix D for details) . Since $u_0=0$, we concentrate on the change in u_1 as W_{ijt} increases. Suppose

$$\frac{\partial\varphi\left(X_{ijt}\right)}{\partial W_{ijt}}=\frac{\partial\varphi_W\left(W_{ijt},T_{ijt},\varsigma_{t+1}^w\right)}{\partial W_{ijt}}\geq 0.$$

Therefore, any future state is increasing in the state variable W_{ijt} in period t, i.e.,

$$\frac{\partial W_{ijt+s}}{\partial W_{ijt}}\geq 0 \text{ for any } s\geq 1.$$

Then from

$$\frac{\partial u_1\left(X_{ijt+s}\right)}{\partial W_{ijt}}=\frac{\partial u_1\left(X_{ijt+s}\right)}{\partial W_{ijt+s}}\times\frac{\partial W_{ijt+s}}{\partial W_{ijt}}$$

it follows that $\frac{\partial u_1\left(X_{ijt+s}\right)}{\partial W_{ijt}}\geq 0$ if and only if $\frac{\partial u_1\left(X_{ijt+s}\right)}{\partial W_{ijt+s}}>0$.

Let us recap and summarize the implications of the above under the assumption $\frac{\partial\varphi}{\partial x}\geq 0$ for some state variable $x\in\{W_{ijt},I_{ijt},H_{ijt},D_{ijt}\}$. If $\frac{\partial u_1}{\partial x}<0$, then the deterministic utility u_1 *in any future state* is decreasing in state variable x. Thus, it is impossible for $v_1^{(1)}$ to be increasing in x, i.e., $\frac{\partial u_1}{\partial x}<0$ implies $\frac{\partial v_1^{(1)}}{\partial x}\leq 0$. Similarly, $\frac{\partial u_1}{\partial x}>0$ implies $\frac{\partial v_1^{(1)}}{\partial x}\geq 0$. Thus,

$$\frac{\partial p_1}{\partial x}=\frac{e^{v_0-v_1}}{\left(1+e^{v_0-v_1}\right)^2}\left(\frac{\partial u_1}{\partial x}+\frac{\partial v_1^{(1)}}{\partial x}\right)>0 \text{ if and only if } \frac{\partial u_1}{\partial x}>0.\quad\square$$

Proof of Proposition 3. Without loss of generality we set $i=1$ and $j=2$. From Proposition 2,

$$\frac{\partial p_1\left(X_t\right)}{\partial X_{it}} = \frac{e^{v_0(X_t)-v_1(X_t)}}{\left(1+e^{v_0(X_t)-v_1(X_t)}\right)^2}\left(\frac{\partial u_1}{\partial X_{it}}+\frac{\partial v_1^{(1)}}{\partial X_{it}}\right),\ i\ \in\{1,2\}, \tag{55}$$

we rewrite linear state transition function as

$$\varphi_i\left(X_{it},t,\zeta_{it}\right)=\sigma_{0i}+\sigma_{1i}X_{it}+\sigma_{2i}t+\zeta_{it+1}. \tag{56}$$

By supposition,

$$\frac{\partial u_1}{\partial X_{1t}}=\theta_1>\theta_2=\frac{\partial u_1}{\partial X_{2t}}>0 \tag{57}$$

$$\frac{\partial \varphi_1}{\partial X_{1t}}=\sigma_{11}>\sigma_{12}=\frac{\partial \varphi_2}{\partial X_{2t}}. \tag{58}$$

We see from (55) and (57) that a proof of Propostion 3 is complete if we can show that

$$\frac{\partial u_1}{\partial X_{1t}}>\frac{\partial u_1}{\partial X_{2t}} \text{ implies } \frac{\partial v_1^{(1)}}{\partial X_{1t}}\geq\frac{\partial v_1^{(1)}}{\partial X_{2t}}. \tag{59}$$

The above holds for $t=T$ because $v_1^{(1)}\left(X_T\right)=v_0^{(1)}\left(X_T\right)=0$. We begin by show that (59) holds for the case of $t=T-1$, i.e., there is one period remaining after the current period. We then show that the result continue to hold for $t<T-1$. For the reminder of the proof, we assume

$$\frac{\partial u_1}{\partial X_{1t}}>\frac{\partial u_1}{\partial X_{2t}}. \tag{60}$$

Suppose that $t=T-1$. Then

$$v_1^{(1)}\left(X_{T-1}\right)=\beta E\left[V\left(\tilde{X}_T\right)|X_{T-1},d_{T-1}=1\right]=\beta E\left[\max\{\tilde{\varepsilon}_{0T},u_1\left(\tilde{X}_t\right)+\tilde{\varepsilon}_{1T}\}|X_{T-1}\right]. \tag{61}$$

Recall that

$$u_1\left(\tilde{X}_T\right)=\theta_{0T}+\theta'\begin{bmatrix}\sigma_{01}+\sigma_{11}X_{1T-1}+\sigma_{21}\left(T-1\right)+\tilde{\zeta}_{1T}\\ \vdots\\ \sigma_{0n}+\sigma_{11}X_{nT-1}+\sigma_{2n}\left(T-1\right)+\tilde{\zeta}_{nT}\\ T\end{bmatrix}. \tag{62}$$

Let $\zeta_t = (\zeta_{1t}, \dots \zeta_{nt})$, $\varepsilon_t = (\varepsilon_{0t}, \varepsilon_{1t})$,

$$\begin{aligned}
&\varphi(X_t, \zeta_{t+1}) = (\varphi_1(X_{1t}, t, \zeta_{it+1}), \dots, \varphi_n(X_{nt}, t, \zeta_{nt+1}), \varphi_{n+1}(t)) \\
&\Omega_0(\zeta, \varepsilon | X_{T-1}) = \{(\zeta, \varepsilon) : \varepsilon_0 \geq u_1(\varphi(X_{T-1}, \zeta)) + \varepsilon_1\} \qquad (63) \\
&\Omega_1(\zeta, \varepsilon | X_{T-1}) = \{(\zeta, \varepsilon) : \varepsilon_0 < u_1(\varphi(X_{T-1}, \zeta)) + \varepsilon_1\}
\end{aligned}$$

i.e., $\Omega_0(\zeta, \varepsilon | X_{T-1})$ is the set of realizations of random variables $\left(\tilde{\zeta}_T, \tilde{\varepsilon}_T\right)$ for which it is optimal to quit in period T, and $\Omega_1(\zeta, \varepsilon | X_{T-1})$ is the set of realizations of random variables $\left(\tilde{\zeta}_T, \tilde{\varepsilon}_T\right)$ for which it is optimal to continue in period T. With this notation, we can rewrite (61) as

$$v_1^{(1)}(X_{T-1}) = \beta \left(\int\limits_{\Omega_0(\zeta, \varepsilon | X_{T-1})} \varepsilon_0 dG(\zeta, \varepsilon) + \int\limits_{\Omega_1(\zeta, \varepsilon | X_{T-1})} (u_1(\varphi(X_{T-1}, \zeta)) + \varepsilon_1) \, dG(\zeta, \varepsilon) \right).$$

Let e_i denote an $n+1$ dimensional vector with the i^{th} element equal to 1 and 0 elsewhere, e.g., $e_2 = (0, 1, 0, \dots, 0)$. Let Δ be a positive value. We will next show that

$$v_1^{(1)}(X_{T-1} + \Delta e_2) - v_1^{(1)}(X_{T-1}) \leq v_1^{(1)}(X_{T-1} + \Delta e_1) - v_1^{(1)}(X_{T-1}) \text{ for any } \Delta > 0$$

and that this inequality implies $\frac{\partial v_1^{(1)}}{\partial X_{2T-1}} \leq \frac{\partial v_1^{(1)}}{\partial X_{1T-1}}$. Note that

$$\begin{aligned}
u_1(\varphi(X_{T-1} + \Delta e_2, \zeta)) =& u_1(\varphi(X_{T-1}, \zeta)) + \left(\frac{\partial u_1}{\partial X_{2T}}\right)\left(\frac{\partial \varphi_2}{\partial X_{2T-1}}\right)\Delta \\
&= u_1(\varphi(X_{T-1}, \zeta)) + \theta_2 \sigma_{12} \Delta \qquad (64) \\
&< u_1(\varphi(X_{T-1}, \zeta)) + \theta_1 \sigma_{11} \Delta \quad (see (57) \; and \; (58)) \\
&= u_1(\varphi(X_{T-1} + \Delta e_1, \zeta)).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&v_1^{(1)}\left(X_{T-1}+\Delta e_2\right)\\
=&\beta\left(\int\limits_{\Omega_0(\zeta,\varepsilon|X_{T-1}+\Delta e_2)}\varepsilon_0 dG\left(\zeta,\varepsilon\right)+\int\limits_{\Omega_1(\zeta,\varepsilon|X_{T-1}+\Delta e_2)}\left(u_1\left(\varphi\left(X_{T-1}+\Delta e_2,\zeta\right)\right)+\varepsilon_1\right)dG\left(\zeta,\varepsilon\right)\right)\\
&\leq\beta\left(\int\limits_{\Omega_0(\zeta,\varepsilon|X_{T-1}+\Delta e_2)}\varepsilon_0 dG\left(\zeta,\varepsilon\right)+\int\limits_{\Omega_1(\zeta,\varepsilon|X_{T-1}+\Delta e_2)}\left(u_1\left(\varphi\left(X_{T-1}+\Delta e_1,\zeta\right)\right)+\varepsilon_1\right)dG\left(\zeta,\varepsilon\right)\right)\\
&\leq\beta\left(\int\limits_{\Omega_0(\zeta,\varepsilon|X_{T-1}+\Delta e_1)}\varepsilon_0 dG\left(\zeta,\varepsilon\right)+\int\limits_{\Omega_1(\zeta,\varepsilon|X_{T-1}+\Delta e_1)}\left(u_1\left(\varphi\left(X_{T-1}+\Delta e_1,\zeta\right)\right)+\varepsilon_1\right)dG\left(\zeta,\varepsilon\right)\right)\\
&=v_1^{(1)}\left(X_{T-1}+\Delta e_1\right).
\end{aligned}$$

The first inequality follows from (64). The second inequality follows from the definition of $\Omega_0\left(\zeta,\varepsilon|X_{T-1}+\Delta e_1\right)$ and $\Omega_1\left(\zeta,\varepsilon|X_{T-1}+\Delta e_1\right)$ that are sets of realizations associated with each optimal decision (quit or continue) for a given state. From

$$v_1^{(1)}\left(X_{T-1}+\Delta e_2\right)\leq v_1^{(1)}\left(X_{T-1}+\Delta e_1\right) \tag{65}$$

for any $\Delta>0$, it follows that

$$\lim_{\Delta\to 0}\frac{v_1^{(1)}(X_{T-1}+\Delta e_2)-v_1^{(1)}(X_{T-1})}{\Delta}=\frac{\partial v_1^{(1)}}{\partial X_{2T-1}}\leq\frac{\partial v_1^{(1)}}{\partial X_{1T-1}}=\lim_{\Delta\to 0}\frac{v_1^{(1)}(X_{T-1}+\Delta e_1)-v_1^{(1)}(X_{T-1})}{\Delta}. \tag{66}$$

Now suppose that $t=T-2$. Note that

$$\varphi\left(X_t+\Delta e_i,\zeta\right)=\varphi\left(X_t,\zeta\right)+\sigma_{1i}\Delta e_i \text{ (see (56) and (63)).} \tag{67}$$

Therefore,

$$
\begin{aligned}
&v_1^{(1)}\left(X_{T-2}+\Delta e_2\right)\\
&\quad=\beta\left(\begin{array}{l}\displaystyle\int_{\Omega_0(\zeta,\varepsilon|X_{T-2}+\Delta e_2)}\left(\varepsilon_0+\beta E\left[\tilde{\varepsilon}_{0T}\right]\right)dG\left(\zeta,\varepsilon\right)\\ \displaystyle+\int_{\Omega_1(\zeta,\varepsilon|X_{T-2}+\Delta e_2)}\left(\begin{array}{l}u_1\left(\varphi\left(X_{T-2}+\Delta e_2,\zeta\right)\right)+\varepsilon_1\\+v_1^{(1)}\left(\varphi\left(X_{T-2}+\Delta e_2,\zeta\right)\right)\end{array}\right)dG\left(\zeta,\varepsilon\right)\end{array}\right)\\
&\quad=\beta\left(\begin{array}{l}\displaystyle\int_{\Omega_0(\zeta,\varepsilon|X_{T-2}+\Delta e_2)}\left(\varepsilon_0+\beta\gamma\right)dG\left(\zeta,\varepsilon\right)\\ \displaystyle+\int_{\Omega_1(\zeta,\varepsilon|X_{T-2}+\Delta e_2)}\left(\begin{array}{l}u_1\left(\varphi\left(X_{T-2}+\Delta e_2,\zeta\right)\right)+\varepsilon_1\\+v_1^{(1)}\left(\varphi\left(X_{T-2},\zeta\right)+\sigma_{12}\Delta e_2\right)\end{array}\right)dG\left(\zeta,\varepsilon\right)\end{array}\right)\quad\text{see (67)}\\
&\quad\leq\beta\left(\begin{array}{l}\displaystyle\int_{\Omega_0(\zeta,\varepsilon|X_{T-2}+\Delta e_2)}\left(\varepsilon_0+\beta\gamma\right)dG\left(\zeta,\varepsilon\right)\\ \displaystyle+\int_{\Omega_1(\zeta,\varepsilon|X_{T-2}+\Delta e_2)}\left(\begin{array}{l}u_1\left(\varphi\left(X_{T-2}+\Delta e_1,\zeta\right)\right)+\varepsilon_1\\+v_1^{(1)}\left(\varphi\left(X_{T-2},\zeta\right)+\sigma_{11}\Delta e_1\right)\end{array}\right)dG\left(\zeta,\varepsilon\right)\end{array}\right)\\
&\quad\leq\beta\left(\begin{array}{l}\displaystyle\int_{\Omega_0(\zeta,\varepsilon|X_{T-2}+\Delta e_1)}\left(\varepsilon_0+\beta\gamma\right)dG\left(\zeta,\varepsilon\right)\\ \displaystyle+\int_{\Omega_1(\zeta,\varepsilon|X_{T-2}+\Delta e_1)}\left(\begin{array}{l}u_1\left(\varphi\left(X_{T-2}+\Delta e_1,\zeta\right)\right)+\varepsilon_1\\+v_1^{(1)}\left(\varphi\left(X_{T-2},\zeta\right)+\sigma_{11}\Delta e_1\right)\end{array}\right)dG\left(\zeta,\varepsilon\right)\end{array}\right)\\
&\quad=v_1^{(1)}\left(X_{T-2}+\Delta e_1\right).
\end{aligned}
\tag{68}
$$

for any $\Delta > 0$. The first inequality follows from (58), (64), and (65). The second inequality follows from replacing the optimal decisions in period $T-1$ conditioned on state $X_{T-2}+\Delta e_2$ in period $T-2$ with optimal decisions conditioned on state $X_{T-2}+\Delta e_1$. Therefore,

$$\lim_{\Delta\to 0}\frac{v_1^{(1)}(X_{T-2}+\Delta e_2)-v_1^{(1)}(X_{T-2})}{\Delta}=\frac{\partial v_1^{(1)}}{\partial X_{2T-2}}\leq\frac{\partial v_1^{(1)}}{\partial X_{1T-2}}=\lim_{\Delta\to 0}\frac{v_1^{(1)}(X_{T-2}+\Delta e_2)-v_1^{(1)}(X_{T-2})}{\Delta}.$$

Thus, by induction, the result holds for $t=T-\tau$ for any $\tau\in\{1,\ldots,T-1\}$. $\square$

Proof of Proposition 4 . We rewrite Equation (27) using our streamlined notation

$$p_1\left(X_t\right)=\frac{1}{\left(1+e^{v_0\left(X_t\right)-v_1\left(X_t\right)}\right)}$$

and implicitly differentiate with respect to state transition parameters σ_i

$$\begin{aligned}\frac{\partial p_1\left(X_t\right)}{\partial\sigma_i}&=\frac{e^{v_0\left(X_t\right)-v_1\left(X_t\right)}}{\left(1+e^{v_0\left(X_t\right)-v_1\left(X_t\right)}\right)^2}\left(\frac{\partial u_1\left(X_t\right)}{\partial\sigma_i}+\frac{\partial v_1^{(1)}\left(X_t\right)}{\partial\sigma_i}\right)\\&=\frac{e^{v_0\left(X_t\right)-v_1\left(X_t\right)}}{\left(1+e^{v_0\left(X_t\right)-v_1\left(X_t\right)}\right)^2}\frac{\partial v_1^{(1)}\left(X_t\right)}{\partial\sigma_i},\end{aligned}\tag{69}$$

(i.e., u_1 drops out because it only depends on the current state, and not the future states that are influenced by σ_i). We define

$$\Omega_0\left(\zeta,\varepsilon|X_t\right)=\left\{(\zeta,\varepsilon):\varepsilon_0\geq u_1\left(\varphi\left(X_t,\zeta\right)\right)+\varepsilon_1+v_1^{(1)}\left(\varphi(X_t,\zeta)\right)\right\}$$

$$\Omega_1\left(\zeta,\varepsilon|X_t\right)=\left\{(\zeta,\varepsilon):\varepsilon_0<u_1\left(\varphi\left(X_t,\zeta\right)\right)+\varepsilon_1+v_1^{(1)}\left(\varphi(X_t,\zeta)\right)\right\}$$

i.e., $\Omega_0\left(\zeta,\varepsilon|X_t\right)$ is the set of realizations of random variables $\left(\tilde{\zeta}_{t+1},\tilde{\varepsilon}_{t+1}\right)$ for which it is optimal to quit in period $t+1$ and $\Omega_1\left(\zeta,\varepsilon|X_t\right)$ is the set of realizations of random variable $\left(\tilde{\zeta}_{t+1},\tilde{\varepsilon}_{t+1}\right)$ for which it is optimal to continue in period $t+1$. Then the conditional valuation function can be expressed as

$$v_1^{(1)}\left(X_t\right)=\beta\left(\begin{array}{l}\displaystyle\int_{\Omega_0(\zeta,\varepsilon|X_t)}\left(\varepsilon_0+\beta E\left[\sum_{s=2}^{T-t-1}\beta^{s-1}\tilde{\varepsilon}_{0t+S}\right]\right)dG\left(\zeta,\varepsilon\right)\\+\displaystyle\int_{\Omega_1(\zeta,\varepsilon|X_t)}\left(u_1\left(\varphi\left(X_t,\zeta\right)\right)+\varepsilon_1+v_1^{(1)}\left(\varphi\left(X_t,\zeta\right)\right)\right)dG\left(\zeta,\varepsilon\right)\end{array}\right).$$

From the definition of $\Omega_0\left(\zeta,\varepsilon|X_t\right)$ and $\Omega_1\left(\zeta,\varepsilon|X_t\right)$, it follows that the two integrands form a piecewise continuous function over realizations of $\left(\tilde{\zeta}_{t+1},\tilde{\varepsilon}_{t+1}\right)$. To clarify this point, we rewrite the above as

$$\beta\left(\int_{\Omega_0(\zeta,\varepsilon|X_t)}h_1\left(\zeta,\varepsilon,\sigma_{ik}\right)dG\left(\zeta,\varepsilon\right)+\int_{\Omega_1(\zeta,\varepsilon|X_t)}h_2\left(\zeta,\varepsilon,\sigma_{ik}\right)dG\left(\zeta,\varepsilon\right)\right),$$

where $h_1\left(\zeta,\varepsilon,\sigma_{ik}\right)=\varepsilon_0+E\left[\sum_{s=2}^{T-t-1}\beta^{s-1}\tilde{\varepsilon}_{0t+S}\right]=u_1\left(\varphi\left(X_t,\zeta\right)\right)+\varepsilon_1+v_1^{(1)}\left(\varphi\left(X_t,\zeta\right)\right)=h_2\left(\zeta,\varepsilon,\sigma_{ik}\right)$ at all $(\zeta,\varepsilon)\in\left\{(\zeta,\varepsilon):\varepsilon_0=u_1\left(\varphi\left(X_t,\zeta\right)\right)+\varepsilon_1+v_1\left(\varphi\left(X_t,\zeta\right)\right)\right\}$. Therefore, the partial derivative of $v_1^{(1)}(X_t)$ with respect to σ_{ik} is obtained purely from the partial derivatives of the integrands, i.e.,

$$\begin{aligned}
v_1^{(1)}(X_t) =& \beta \left(\int\limits_{\Omega_0(\zeta,\varepsilon|X_t)} \frac{\partial}{\partial \sigma_{ik}} \left(\varepsilon_0 + \beta E \left[\sum_{s=2}^{T-t-1} \beta^{s-1} \tilde{\varepsilon}_{0t+S} \right] \right) dG(\zeta,\varepsilon) \right. \\
& \left. + \int\limits_{\Omega_1(\zeta,\varepsilon|X_t)} \frac{\partial}{\partial \sigma_{ik}} \left(u_1(\varphi(X_t,\zeta)) + \varepsilon_1 + v_1^{(1)}(\varphi(X_t,\zeta)) \right) dG(\zeta,\varepsilon) \right) \\
=& \beta \left(\int\limits_{\Omega_1(\zeta,\varepsilon|X_t)} \left(\frac{\partial u_1(\varphi(X_t,\zeta))}{\partial \sigma_{ik}} + \frac{\partial v_1^{(1)}(\varphi(X_t,\zeta))}{\partial \sigma_{ik}} \right) dG(\zeta,\varepsilon) \right).
\end{aligned}$$

Note that

$$\begin{aligned}
\frac{\partial u_1(\varphi(X_t,\zeta))}{\partial \sigma_{ik}} &= \frac{\partial u_1(X_{t+1})}{\partial X_{it+1}} \frac{\varphi(X_t,\zeta)}{\partial \sigma_{ik}} \\
\frac{\partial v_1^{(1)}(\varphi(X_t,\zeta))}{\partial \sigma_{ik}} &= \frac{\partial v_1^{(1)}(X_{t+1})}{\partial X_{it+1}} \frac{\varphi(X_t,\zeta)}{\partial \sigma_{ik}}.
\end{aligned}$$

Substituting the above into Equation (69)

$$\frac{\partial p_1(X_t)}{\partial \sigma_{ik}} = \frac{e^{v_0(X_t)-v_1(X_t)}}{\left(1+e^{v_0(X_t)-v_1(X_t)}\right)^2} \beta \left(\int\limits_{\Omega_1(\zeta,\varepsilon|X_t)} \left(\frac{\partial u_1}{\partial X_{it+1}} + \frac{\partial v_1^{(1)}}{\partial X_{it+1}} \right) \frac{\partial \varphi(X_t,\zeta)}{\partial \sigma_{ik}} dG(\zeta,\varepsilon) \right).$$

Finally, as shown in the proof of Proposition 2

$$\frac{\partial u_1(X_{t+1})}{\partial X_{it+1}} > 0 \text{ implies } \frac{\partial v_1^{(1)}(X_{t+1})}{\partial X_{it+1}} \geq 0$$

$$\frac{\partial u_1(X_{t+1})}{\partial X_{it+1}} < 0 \text{ implies } \frac{\partial v_1^{(1)}(X_{t+1})}{\partial X_{it+1}} \leq 0.$$

Therefore, if $\frac{\partial \varphi_i}{\partial \sigma_{ik}} \geq 0$, then the sign of $\frac{\partial p_1(X_t)}{\partial \sigma_{ik}}$ matches the sign of $\frac{\partial u_1}{\partial X_{it+1}}$. $\square$

APPENDIX I

PROPENSITY SCORE WEIGHTING

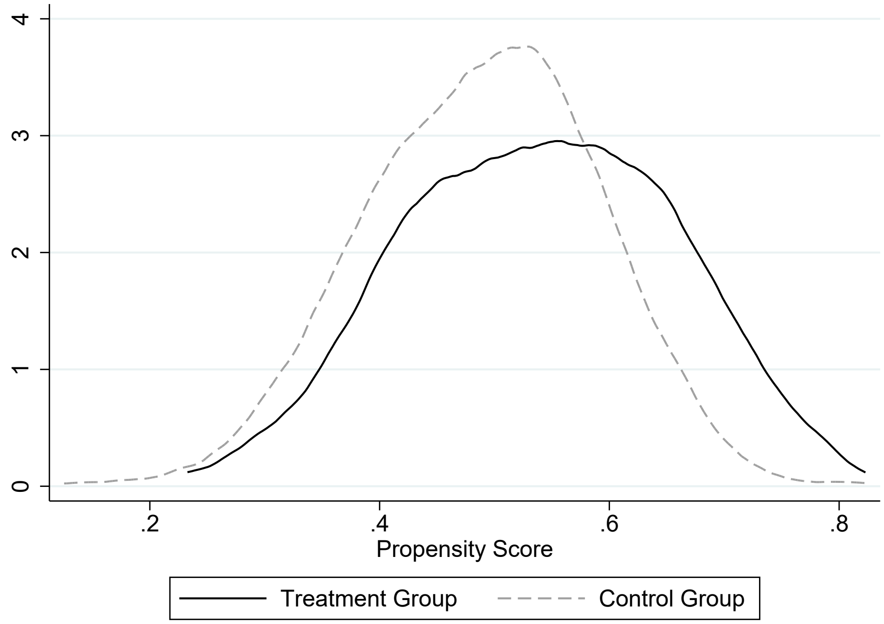
Our use of the inverse propensity score weighting (PSW) procedure follows Abadie (2005). A key advantage of this procedure is that it allows us to utilize the full sample without being restricted to an equal number of treated and control ZIP code areas. This procedure also improves comparability between the treatment and control groups by assigning more weight to ZIP code areas in a group that are more similar to those in the other group. In particular, the weight assigned to each ZIP code area i is:

$$Weight_i = treat_i \frac{p}{p(X_i)} + (1 - treat_i) \frac{1-p}{1-p(X_i)}, \tag{70}$$

where $treat_i$ is a binary variable equal to 1 if the ZIP code area experienced Amazon's integration and 0 otherwise. p is the unconditional probability of belonging to the treatment group while $p(X_i)$ is the probability conditional on the vector of individual ZIP code area attributes X_i. [37] As shown in Figure 20, the estimated propensity scores across the treatment and control groups overlap the most in the middle of the distribution, meaning these ZIP code areas are very similar in terms of their observed characteristics. The PSW procedure we used increases the weight of the observations of these ZIP code areas in the analysis relative to those in the left tail.

[37] We dropped 6 ZIP code areas for which the demographic covariates were missing.

Figure 20: Propensity score weighting



The figure shows the kernel density of the estimated propensity score across the treatment and control groups. It is calculated by using estimates from the likelihood of a ZIP code area belonging to the treatment group.

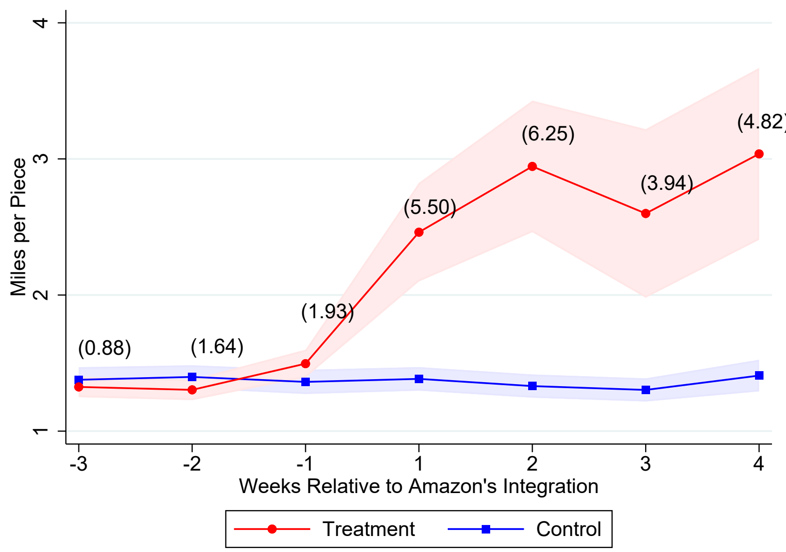
Table 31: Propensity score covariate balance test

	Unweighted Sample			Propensity Score Weighted		
	Estimate	Std. Err.	t-statistic	Estimate	Std. Err.	t-statistic
Mileage per Piece	-0.130	0.058	-2.254	-0.005	0.036	-0.127
Remoteness (Miles)	-0.762	0.491	-1.552	-0.069	0.289	-0.240
Proportion of Fast Deliveries	-0.065	0.011	-5.940	-0.002	0.007	-0.325
Number of Households	447.539	372.845	1.200	21.648	215.484	0.100
Median Annual Household Income	4,137.847	1,749.839	2.365	-199.955	1,036.694	-0.193
Number of Retail Establishments	7.178	4.621	1.553	-0.263	2.712	-0.097
Number of Accommodation and Food Services	7.368	3.482	2.116	-0.003	2.029	-0.002

To validate the PSW approach, we perform a balance test to evaluate whether the covariates are statistically different between the treatment and control ZIP code areas. We regress each covariate in the vector of X_i as a dependent variable on the treatment group indicator $treat_i$, using $Weight_i$ as the regression weight. The idea is that if the means of a covariate are equal across the treatment and control ZIP code areas, the estimate of $treat_i$ should be statistically non-significant. We report the estimate of $treat_i$ for each covariate in the three right-most columns in Table 31. As a comparison, we perform the same regression without weighting the sample and report the estimates on the left side of the table. The estimates of the covariates are all statistically significant when using the unweighted approach (as shown in the left columns), whereas the estimates obtained using weights are nonsignificant at a 0.05 level. This indicates that weighting observations based on $Weight_i$ in Equation (70) indeed eliminates the differences in average characteristics between the treatment and control ZIP code areas.

Figure 21 and Tables 32-34 summarize the results based on the PSW method for the analyses included in Section 4 of the paper in Figure 11 and Tables 14-16. As shown in Tables 32-34, the estimates for the coefficients are consistent with those reported in Tables 14-16 in the main paper.

Figure 21: Effects of Amazon's integration on route density-propensity score weighted



The absolute values of the t-statistics comparing route densities between treatment and control ZIP code areas for each week are reported in parentheses.

Table 32: Relative time model of the effects of Amazon's integration on route density-propensity score weighted

	(1)		(2)	
	Estimate	Std. Err.	Estimate	Std. Err.
3^{rd} Week before Amazon Integration (γ_{-3})	-0.235	(0.160)	-0.247	(0.160)
2^{nd} Week before Amazon Integration (γ_{-2})	-0.086	(0.163)	-0.109	(0.164)
1^{st} Week after Amazon Integration (γ_1)	1.110***	(0.179)	1.201***	(0.189)
2^{nd} Week after Amazon Integration (γ_2)	1.402***	(0.194)	1.503***	(0.205)
3^{rd} Week after Amazon Integration (γ_3)	1.180***	(0.196)	1.281***	(0.206)
4^{th} Week after Amazon Integration (γ_4)	1.468***	(0.202)	1.550***	(0.209)
Observations	7,735		7,735	
R-squared	0.396		0.397	
Week FE	Yes		Yes	
Service area FE	Yes		Yes	
Control ($S_{i,t}$)	No		Yes	

Notes: 1. Robust standard errors are clustered at ZIP code area level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specification in Column (2) includes as a predictor the weekly number of stops in each ZIP code area (S_{it}).

Table 33: Average effects of Amazon's integration on route density-propensity score weighted

	(1)		(2)	
	Estimate	Std. Err.	Estimate	Std. Err.
Integration × Treat (δ)	1.352***	(0.283)	1.449***	(0.362)
Observations	7,735		7,735	
R-squared	0.396		0.396	
Week FE	Yes		Yes	
Service area FE	Yes		Yes	
Control ($S_{i,t}$)	No		Yes	

Notes: 1. Robust standard errors are clustered at ZIP code area level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specification in Column (2) includes as a predictor the weekly number of stops in each ZIP code area (S_{it}).

Table 34: Heterogeneity effects of Amazon's integration on route density-propensity score weighted

	(1)	(2)	(3)	(4)	(5)	(6)
Integration × Treat (δ)	0.223	0.275	-0.590	-0.505	-1.996**	-1.953**
	(0.553)	(0.561)	(0.380)	(0.325)	(0.890)	(0.843)
Integration × Treat						
× Remoteness	0.068*	0.068*			0.080**	0.079**
	(0.040)	(0.039)			(0.038)	(0.038)
Integration × Treat						
× Proportion of Fast Deliveries			7.768***	7.810***	8.053***	8.068***
			(2.334)	(2.364)	(2.400)	(2.424)
Remoteness (Miles)	0.019	0.019			0.025	0.025
	(0.066)	(0.066)			(0.058)	(0.058)
Proportion of Fast Deliveries			-0.573	-0.629	-0.780	-0.801
			(0.688)	(0.705)	(0.689)	(0.709)
Observations	7,735	7,735	7,727	7,727	7,735	7,735
R-squared	0.402	0.402	0.461	0.461	0.469	0.469
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Service area FE	Yes	Yes	Yes	Yes	Yes	Yes
Control ($S_{i,t}$)	No	Yes	No	Yes	No	Yes

Notes: 1. Robust standard errors are clustered at ZIP code area level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specifications in Columns (2), (4), and (6) include as a predictor the weekly number of stops in each ZIP code area (S_{it}).

APPENDIX J

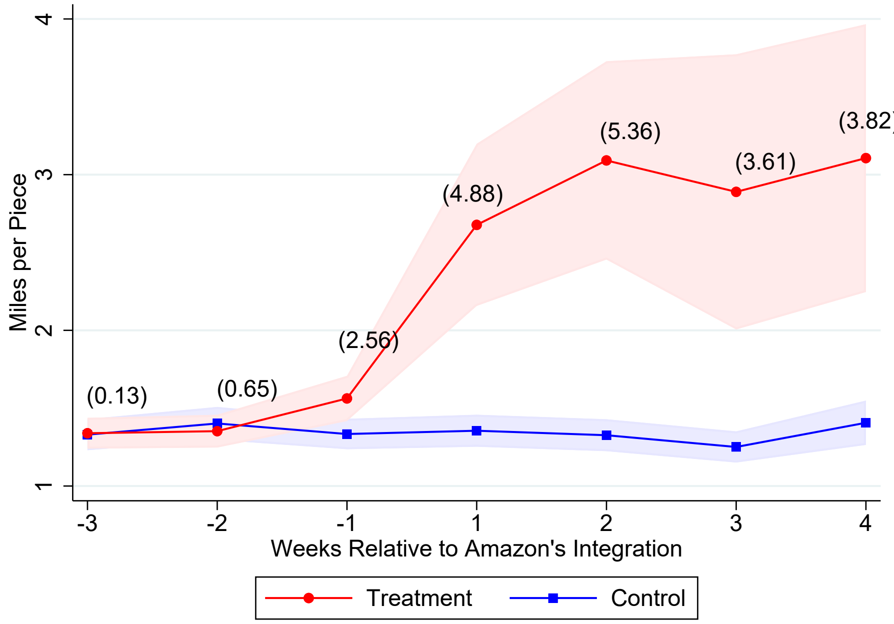
PROPENSITY SCORE MATCHING

Rather than using regression weights to eliminate the differences in the covariates between the treatment and control ZIP code areas, the propensity score matching (PSM) procedure uses a one-to-one matching without replacement between the ZIP code areas. To implement this procedure, we follow Son *et al.* (2019). Specifically, we employ, for the matching procedure, a standard caliper size which is 0.2 times the standard deviation of the propensity scores. This yields a subsample consisting of 300 ZIP code areas in the treatment group and 300 matching ZIP code areas in the control group. We verify our matching performance by using a t-test to compare the group means of the covariates after matching and report the results in Table 35. The values of the t-statistics confirm that, after matching, the differences between the group means for the covariates are not statistically different from zero at the 0.05 level of significance. We then use the matched ZIP code areas to plot the differences in the weekly route density in Figure 22. As shown in this figure, the differences in the weekly average measures of route density are consistently significant only after the treatment application. We also repeat the analyses to estimate the coefficients, as in Tables 14-16 in the paper. As shown in Tables 36-38, the coefficients estimated using the PSM procedure are consistent with those reported in Tables 14-16 in the paper.

Table 35: Covariate balance test after propensity score matching

	Treatment Group Mean	Control Group Mean	t-statistic (Diff)
Mileage per Piece	1.418	1.355	-0.875
Remoteness (Miles)	16.541	16.712	0.244
Proportion of Fast Deliveries	0.379	0.358	-1.404
Number of Households	12,068.437	11,950.087	-0.237
Median Annual Household Income	67,655.893	68,363.603	0.300
Number of Retail Establishments	98.057	98.213	0.026
Number of Accommodation and Food Services	74.793	76.047	0.268

Figure 22: Effects of Amazon's integration on route density after propensity score matching



The absolute values of the t-statistics comparing route densities between treatment and control ZIP code areas for each week are reported in parentheses.

Table 36: Relative time model of the effects of Amazon's integration on route density after propensity score Matching

	(1)		(2)	
	Estimate	Std. Err.	Estimate	Std. Err.
3^{rd} Week before Amazon Integration (γ_{-3})	0.070	(0.192)	0.066	(0.184)
2^{nd} Week before Amazon Integration (γ_{-2})	0.226	(0.226)	0.219	(0.205)
1^{st} Week after Amazon Integration (γ_1)	1.440***	(0.386)	1.464***	(0.482)
2^{nd} Week after Amazon Integration (γ_2)	1.792***	(0.435)	1.818***	(0.523)
3^{rd} Week after Amazon Integration (γ_3)	1.612***	(0.564)	1.638**	(0.670)
4^{th} Week after Amazon Integration (γ_4)	1.802***	(0.513)	1.824***	(0.602)
Observations	4,200		4,200	
R-squared	0.360		0.360	
Week FE	Yes		Yes	
Service area FE	Yes		Yes	
Control ($S_{i,t}$)	No		Yes	

Notes: 1. Robust standard errors are clustered at ZIP code area level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specification in Column (2) includes as a predictor the weekly number of stops in each ZIP code area (S_{it}).

Table 37: Average effects of Amazon's integration on route density after propensity score matching

	(1)		(2)	
	Estimate	Std. Err.	Estimate	Std. Err.
Integration $\times$ Treat (δ)	1.561***	(0.345)	1.598***	(0.471)
Observations	4,200		4,200	
R-squared	0.360		0.360	
Week FE	Yes		Yes	
Service area FE	Yes		Yes	
Control ($S_{i,t}$)	No		Yes	

Notes: 1. Robust standard errors are clustered at ZIP code area level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specification in Column (2) includes as a predictor the weekly number of stops in each ZIP code area (S_{it})

Table 38: Heterogeneity effects of Amazon's integration on route density after propensity score matching

	(1)	(2)	(3)	(4)	(5)	(6)
Integration $\times$ Treat (δ)	0.192	0.089	-0.633	-0.597	-2.905*	-3.004*
	(0.843)	(0.846)	(0.632)	(0.540)	(1.668)	(1.592)
Integration $\times$ Treat						
$\times$ Remoteness	0.086	0.088			0.124*	0.125*
	(0.061)	(0.060)			(0.067)	(0.066)
Integration $\times$ Treat						
$\times$ Proportion of Fast Deliveries			7.915**	7.935**	8.796***	8.767**
			(3.083)	(3.139)	(3.399)	(3.431)
Remoteness (Miles)	0.089	0.089			0.064	0.064
	(0.077)	(0.077)			(0.070)	(0.070)
Proportion of Fast Deliveries			-1.118	-1.144	-1.845	-1.802
			(1.029)	(1.072)	(1.155)	(1.196)
Observations	4,200	4,200	4,200	4,200	4,200	4,200
R-squared	0.375	0.375	0.418	0.418	0.438	0.438
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Service area FE	Yes	Yes	Yes	Yes	Yes	Yes
Control ($S_{i,t}$)	No	Yes	No	Yes	No	Yes

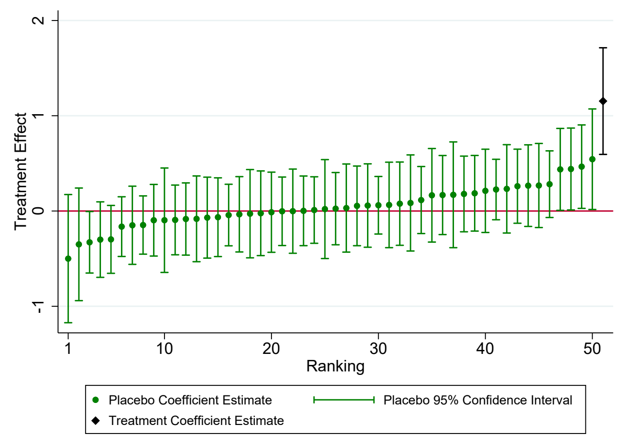
Notes: 1. Robust standard errors are clustered at ZIP code area level.
2. *, ** and *** denote statistical significance at 10 percent, 5 percent and 1 percent levels, respectively.
3. Specifications in Columns (2), (4), and (6) include as a predictor the weekly number of stops in each ZIP code area (S_{it}).

APPENDIX K

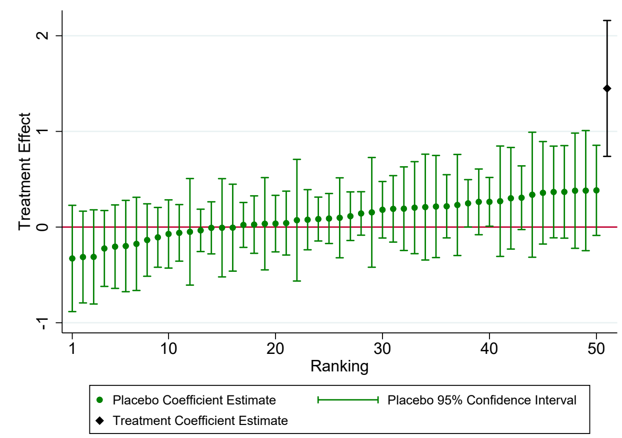
RANDOM IMPLEMENTATION TESTS

Figure 23: Random implementation test of average treatment effects

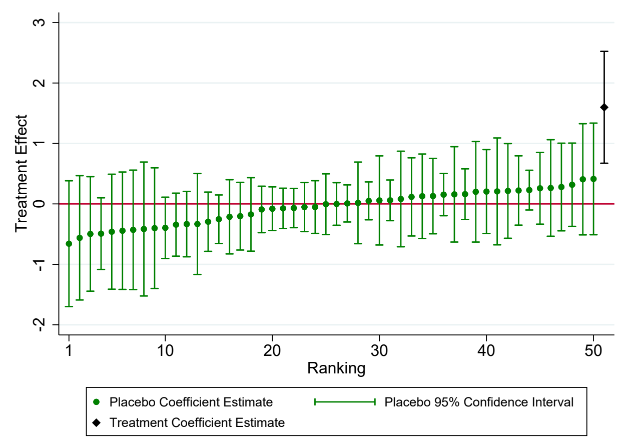
(A) With no propensity score analysis



(B) Using propensity score weighting



(C) Using propensity score matching



Placebo intervention dates and treatment were randomly assigned 50 times. Each point estimate and its 95% confidence intervals correspond to an estimation result of the average treatment effect using Equation 32. The coefficient and confidence interval for the true data are represented on the far right (ranking 51).