

Squeezing more Juice out of Lime: A Novel High-dimensional Pricing Algorithm *

Richard Faltings[†]

March 17, 2024

Abstract

Sophisticated pricing algorithms used by digital transportation platforms have renewed interest in price control policies, but little evidence exists on their redistributive effects. This paper studies a uniform price mandate in the market for shared electric vehicle platforms in Washington, D.C., which prohibits origin- and destination-based pricing. To compute price equilibria encompassing hundreds of prices for specific origin-destination pairs, I develop a new simulation-based pricing algorithm, adapted from the reinforcement learning literature. I apply the algorithm to a demand system estimated using geolocation data from all firms in the market. In the counterfactual exercise, I find that the redistributive effects of the price controls are mild, and mainly serve riders in the periphery of the city. Furthermore, I find that relaxing the price controls increases rides taken by consumers by 41%, firm profits by 34%, and increases consumer welfare by more than double the profit increase (80% of firm profits).

Keywords: Price discrimination, Transportation pricing regulation, Platform competition, Redistribution and efficiency, Algorithmic pricing

JEL Codes: L11, R41, D63, C63

*I would like to thank Eugenio Miravete, Daniel Akerberg, Jorge Balat, Caroline Thomas, Robert Town, and Victoria Marone for their continued advice and support. I would also like to thank Milena Almagro, Volker Nocke, Julian Wright, Martin Peitz, Lauri Kytömaa, Justus Timmers, Ignacio Núñez as well as other participants at the UT Austin IO workshop and the MaCCI Summer School in Platform Economics for valuable comments on this work. I acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper.

[†]The University of Texas at Austin: richard.faltings@utexas.edu

1 Introduction

Price controls are generally agreed to reduce overall economic welfare as they obscure the true economic costs and benefits of an individual’s decisions (Hayek, 1945). Nevertheless, they are often used as a policy tool to address concerns for various kinds of fairness. Controls on the overall level of prices are generally meant to correct for excessive market power, as is the case with the price gouging laws that were regularly invoked in the wake of the COVID-19 pandemic, or to curb runaway wartime inflation (Galbraith, 1980; Taussig, 1919). Controls on the *flexibility* of prices (over time, or over consumer characteristics) can instead serve as a form of cross-subsidization between different groups of consumers.¹ For example, the Affordable Care Act’s ban on price discrimination against individuals with pre-existing conditions effectively subsidizes less healthy individuals. This type of redistributive policy can be justified when imperfect information hinders non-distortionary transfers in the spirit of the Second Theorem of Welfare Economics (Arrow, 1951). In such a second-best scenario, price controls are an attractive policy tool insofar as they target redistribution effectively and do so at a minimal cost to economic efficiency.²

This paper investigates one such price control policy in the context of dockless shared vehicle platforms in Washington D.C. The aims of the paper are to determine: (i) whether the redistributive effects of transportation price controls are progressive in terms of incomes and (ii) the overall cost of the redistribution in terms of foregone consumer surplus and industry profits. In order to credibly do so, this paper introduces a novel simulation-based stochastic gradient descent pricing algorithm derived from the reinforcement learning literature. This allows me to compute highly flexible counterfactual pricing equilibria with hundreds of origin- and destination-specific prices per firm while accounting for the stochastic nature of demand and complex capacity constraints. This method can also flexibly incorporate restrictions on the price schedule by specifying a functional form, as the gradient can simply be passed through via the chain-rule.³

In the context of transportation, price controls have a long history stretching back to the Interstate Commerce Act of 1887 which regulated prices in the railroad industry and es-

¹Controls on pricing flexibility thus cover both third-degree price discrimination (on the basis of different elasticities), and firms setting prices differently according to differences in the marginal cost of serving different consumers.

²Since policies limiting price discrimination will lead to higher prices for some consumers and lower prices for others, they can be considered under the umbrella of second-best redistributive taxation, as considered by Mirrlees (1971). See Boadway and Keen (2000) for a comprehensive review of this subject.

³The vehicles are electric scooters or bicycles. The term ‘dockless’ contrasts these platforms with station-based bikeshare systems.

established the first federal regulatory agency. A notable feature of the regulations was that short-haul rates had to match longer haul rates that traveled the same path. Though the political motivation of the regulation was debated for over a century, the regulation *de facto* subsidized firms with demand for short-haul shipping, which were mostly in the agricultural sector (Gilligan, Marshall, and Weingast, 1990). In the urban transportation context studied in this paper, many public transit systems do not price on the basis of marginal costs, thus effectively implementing a type of price control. In major cities such as New York and Paris, among others, subway and bus trips are all priced the same, regardless of trip length, origin, or destination. Much like the railroad price controls, this has the effect of subsidizing travel to farther peripheral areas, mimicking direct policies such as tax-deductible commuter benefits. A possible rationale for this type of redistribution is that lower-income individuals sort themselves into more distant, cheaper neighborhoods. Though plausible in simple theoretical models, the real-world sorting patterns of a city are complex and can vary according to the transportation technologies available at the time of a city’s development (Anas, Arnott, and Small, 1998). The desirability of redistribution from the central to the peripheral areas of a city is thus ambiguous.

The question of transportation price controls has appeared in renewed form with a tide of privately owned online transportation platforms (such as Uber and Lyft) which have shown the potential for more sophisticated pricing systems to deliver large gains in economic efficiency (see for example, Castillo, 2022 on surge pricing). However, the complex nature of these pricing systems has raised questions on their redistributive effects.⁴ In response, policymakers have proposed and sometimes implemented price controls, despite lacking a clear picture on what kind of redistribution such a policy would actually effect, as was the case in the market for shared vehicle platforms in Washington, D.C.

To study these price controls, I first collect real-time vehicle location data from the Application Programming Interfaces (APIs) of all active firms in the market, which were made public by a regulatory mandate.⁵ The APIs were queried at a minute-level frequency in the period of December 2019 to February 2020, preceding the disruptions of the COVID-19 pandemic. The data for each firm consists of real-time longitude and latitude information on available vehicles. Some firms include a persistent identifier for each vehicle, allowing me to track their movement and construct precise trip information.

⁴See Pandey and Caliskan (2021) for an example of price disparities in ridehail across racial and socio-economic divides.

⁵Like an ordinary website, APIs can be accessed via a URL but rather than display a human-readable page, they serve data that is meant to be read by an application.

I build a rich demand model featuring stochastic arrivals of consumers who then choose among the vehicles of the various firms in the market, taking into account their current prices and the expected distance they would need to walk to reach a firm’s nearest vehicle, which is directly informed by the geolocation data. Unlike other recent work in urban transportation, consumers in my model do not face search frictions. This modeling choice is motivated by the ease of finding vehicles through each firm’s application. The model flexibly estimates the values of trips for each origin-destination pair among a discrete set of 22 areas and additionally distinguishes between 4 periods in the day, to account for time-varying demand patterns such as commuting.

Estimation of the model leverages the rigidity of the two-part tariff used by all firms in the status quo to credibly pin down consumers’ price elasticities without the need for an instrument. The estimator makes full use of the richness of the data, using both re-constructed information on trips for certain firms and direct data on the evolution of vehicle availabilities for other firms.

My estimates imply higher price elasticities of demand than some other recent works studying ridehail platforms in the United States, such as Rosaia (2020) and Cohen et al. (2016), but are lower than the elasticities in Buchholz et al. (2020), who study an auction-based ridehail platform. I also find a high disutility of walking to the nearest vehicle, which, when converted to a value of time, is in line with the estimates of Rosaia (2020). The high price elasticities and the importance of vehicle density both increase the value of more sophisticated prices in this market, the former by increasing the benefit of third degree price discrimination, and the latter by increasing the value of balancing vehicle supply across locations.

In order to assess the effects of the regulation, I develop a novel stochastic gradient ascent algorithm derived from policy gradient methods used in the reinforcement learning literature, which can be used to solve for pricing equilibria under various restrictions on the pricing schedule.⁶ Other recent work in economics has focused on Q-learning algorithms from this literature, but this generally requires iterating action-specific values, using a grid to approximate a continuous action space. When the action space is high-dimensional (such as with a complex price schedule), such methods quickly become computationally intractable. The gradient-based method proposed in this paper instead iterates on the optimal prices directly, allowing it to solve for an equilibrium involving hundreds of different prices for each of the eight firms in the market. It can also naturally accommodate a variety of parametric restrictions on the price schedule by simply applying the chain rule.

⁶See Sutton and Barto, 2018 for an overview of policy gradient reinforcement learning algorithms.

The counterfactual exercise consists of comparing simulated equilibria under optimal two-part tariffs (which are compliant with the price control regulations) against equilibria with fully flexible prices, which can vary by trip origin, destination, time of day, and vehicle availability. I find that allowing fully flexible prices yields a substantial increase of 34% in firm profits, or \$6200 per day, but that the increase in consumer surplus is more than double the profit increase (\$14711 per day, equivalent to 80% of firm profits in the two-part tariff equilibrium), implying that the price control policy imposes a heavy cost on market efficiency. Furthermore, I find that the redistributive effects of the policy are not progressive, but rather inverse U-shaped, such that they mainly benefit the middle of the income distribution. This occurs because the price controls most directly benefit riders taking trips away from the city center and the income distribution varies non-monotonically with distance from the city center. These results thus cast doubt on the effectiveness of these price controls in promoting access to transportation for lower income individuals.

1.1 Related Work

This paper contributes to a recent literature on the efficiency of different pricing schemes in transportation markets, starting with Buchholz (2021), who demonstrated that the simple tariff structure used by taxicabs in New York produces a significant spatial misallocation of capacity. In counterfactual simulations, more sophisticated tariff structures which better reflected drivers' opportunity costs generated substantial efficiency gains comparable to eliminating matching frictions entirely. Following a similar modeling framework, Braccaccio, Kalouptsi, Papageorgiou, and Rosaia (2020) instead focus on frictions attributable to the search and bargaining process in decentralized markets, suggesting a role for a centralized platform to improve match efficiency, or for taxes to correct the mispricing that occurs through bargaining. Rosaia (2020) examines the ridehail market in New York City, focusing on the equilibrium between the two major platforms and the impact of a congestion tariff in relieving the overconcentration of drivers in the busy downtown area.

This work contributes to this literature by explicitly considering the policymaker's redistributive motive for constraining firms to inefficient prices. The empirical setting also differs significantly as it focuses on shared vehicle platforms rather than two-sided markets of drivers and riders. The importance of search frictions is thus vastly diminished, as each platform simply places a fleet of homogenous vehicles at customers' disposal, which can be easily located on a map. On the other hand, since vehicles are rarely relocated by anyone but the customers themselves, the degree of spatial misallocation of vehicles caused by price

distortions is aggravated. The modeling framework in this work can thus be extended to many other transportation markets exhibiting this feature, such as shipping containers⁷, truck trailers⁸, and one-way car rentals.

It is also worth noting that my modeling framework and computational algorithm used in this paper are flexible enough to allow for a form of surge pricing (by allowing a firm’s prices to be conditioned on its current quantity of available vehicles in an area) as in Castillo (2022), but my counterfactual simulations suggest its role is more limited in the empirical context studied here, which can be explained by two major differences. Firstly, the supply of vehicles relative to the quantity of demand is much larger than in rideshare systems, so the stock of vehicles at any one location is depleted more slowly. Secondly, the vehicles themselves cannot respond to price incentives in the same way that rideshare drivers can, so the effects of surge pricing will only take place on the demand side.

This work is also connected to the large literature comparing classical third degree price discrimination with uniform pricing. Theoretical results on price discrimination in oligopolistic contexts, such as by Holmes (1989) and Corts (1998), are ambiguous, as they depend on how competition is intensified or relaxed in the resulting submarkets. Despite this theoretical ambiguity, few papers have examined this question empirically, with a recent example being Adams and Williams (2019) in a retail context. They find an overall benefit to consumers for finer pricing, except for those in rural markets where stores are further apart and competition is weaker. This paper addresses the question of limiting pricing flexibility in a transportation context, where any such controls not only limit third-degree price discrimination, but also any price variation which reflects differences in shadow costs. I find qualitatively similar results, showing a large increase in consumer welfare from more sophisticated pricing, with the only consumers harmed being those in the more rural areas with less vehicles available. In the context of a regulated monopoly, Miravete, Seim, and Thurk (2020) find that the uniformity of alcohol taxes in Pennsylvania effectively acts as a form of progressive redistribution to lower income and less educated individuals. This paper arrives at the opposite conclusion in the urban transportation context, showing that uniformity of transportation prices are not clearly progressive but instead subsidize more distant areas,

⁷Lu, C.-Y. Lee, and L.-H. Lee (2020) study the joint pricing and relocation and pricing problem for containers, but are restrict their setting to two locations and abstract away from the stochastic nature of the problem. In contrast, the reinforcement learning approach in this paper can deal with the larger state spaces that come with many locations and a stochastic problem.

⁸A growing segment of the trucking market operates on the “power only” model, in which a driver picks up an already loaded trailer provided by a third-party. The firms providing such trailers must constantly rebalance their fleets, preferably through appropriate prices. See: <https://convoy.com/blog/drop-and-hook-combinatorial-optimization/>

which are not always higher or lower income than the central part of a city. Finally, several studies investigate the puzzling uniformity of pricing schemes found in many industries, such as DellaVigna and Gentzkow (2019), Cho and Rust (2010), Shiller and Waldfogel (2011), and Orbach and Einav (2007). In general, these papers point to constraints on managerial resources as an explanation for the price uniformity. In contrast, this paper focuses on government mandated price uniformity, and thus bears more direct policy relevance.

Finally, this work also contributes to the understanding of modern reinforcement-learning algorithms in economic research. On the methodological front, such algorithms have recently shown success in dynamic problems with a very large state or action space (Mnih et al., 2015; Silver et al., 2016), where the methods conventionally used in economics would face a curse of dimensionality (see Igami, 2020 and Iskhakov, Rust, and Schjerning, 2020 for a comparison of the approaches). Two major contributors to these successes are the use of simulations to iterate on value and policy functions, and the use of deep neural networks as flexible function approximators. Iteration by simulation has most prominently been used in economics by Pakes and McGuire (2001), but has seen little use in other applied research. Neural networks have recently been used to solve large dynamic macroeconomic models (Azinovic, Gaegauf, and Scheidegger, 2022; Kahou, Fernández-Villaverde, Perla, and Sood, 2021). This paper is the first to propose a tractable framework to apply these techniques to high-dimensional pricing problems. To deal with the continuous action space in such problems, it adapts gradient based methods to iterate on the state-dependent price policy directly via stochastic gradient descent, allowing it to solve for nearly 4,000 equilibrium prices (for each origin, destination, and firm combination) as a function of a high-dimensional state variable. Such a computation would be infeasible with a grid-based approach, as used by Buchholz (2021), as the number of grid points to evaluate would rapidly explode with the number of prices.⁹ The framework proposed here can thus significantly expand the scope of pricing problems that can be studied empirically, without needing to impose restrictions on the number of firms considered or the level of sophistication of their pricing.

Reinforcement learning algorithms have also generated theoretical interest for their potential to reach a supra-competitive pricing equilibrium (Calvano, Calzolari, Denicolo, and Pastorello, 2020). Asker, Fershtman, and Pakes (2023) show that equilibrium behavior depends crucially on the details of how the reinforcement learnings update their prices. However, this literature has only considered Q-learning algorithms, which rely on a discretization of the space of prices, and can mistakenly converge if the algorithm has had insufficient exper-

⁹For a system of K prices and N possible values for each prices, one would have to evaluate a grid of size K^N .

imentation with certain prices. This is again a problem when considering a large schedule of prices, since the algorithm would need to iterate over a large grid. In contrast, the continuous gradient based method presented in this paper iterates on prices directly and thus scales linearly in the number of prices, and its use of stochastic gradient descent naturally seeks to satisfy the first order conditions of a Walrasian equilibrium are fulfilled.¹⁰ This paper thus further reinforces the notion that not all reinforcement learning algorithms are likely to lead to anti-competitive outcomes. As demonstrated in this paper, an alternative rationale for firms’ adoption of such algorithms is their computational advantage in solving for sophisticated price schedules.

The remainder of the work is structured as follows. Section 2 introduces the pricing algorithm used to compute the high-dimensional counterfactual price equilibria. Section 3 introduces the empirical setting and the data used for the analysis. Section 4 explains the economic model. Section 5 details the estimation methodology and resulting estimates. Section 6 introduces the stochastic gradient ascent algorithm used to compute the counterfactual prices and analyzes the counterfactual results. Section 7 concludes. The accompanying online appendix provides additional details on the implementation of the algorithm in Section A.3.

2 Pricing Algorithm

The general application for the pricing algorithm is a multi-product firm that faces (i) stochastic demand, (ii) capacity constraints which can bind across product categories. This describes a wide-variety of industries, such as airlines, hotels, and car rentals, in addition to this paper’s empirical application to scooter platforms. For example, two separate flight itineraries may share capacity on a common leg, and overlapping hotel stays can share capacity on common nights. For car rentals and scooter platforms, the same vehicle could be used for trips to different destinations. In general, these shared capacity constraints create additional linkages in the optimal pricing problem across products, in addition to substitution effects. The consequences are two-fold: the global pricing problem cannot easily be separated into lower-dimensional problems, and the capacity constraints and stochastic demand create complicated non-linearities in the objective function. Exact solutions based on traditional dynamic programming methods are thus infeasible, as solving over the whole state-space will introduce a curse of dimensionality, and a steady-state approximation will

¹⁰Since the conditions for the invertibility of a demand system also guarantee the uniqueness of the Walrasian equilibrium (Berry, Gandhi, and Haile, 2013), the first order conditions will be sufficient in most economic models under empirical study.

not properly capture the effects of the sometimes binding capacity constraints.

To solve these issues, I introduce a simulation-based stochastic gradient descent algorithm. The use of simulation flexibly captures the interaction of capacity constraints and stochastic demand, while the use of gradients significantly improves the rate of convergence in high-dimensional problems.

We begin with the generalized firm pricing problem:

$$\max_{p_i(Q_i, h)} \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{q_i, h | p_i} \left[\mathbb{1}[h_t \neq \bar{h}] \sum_{\ell \in \mathcal{L}} \sum_{m \in \mathcal{M}} (p_{\ell, m, \theta_i}(Q_{f, t}, h_t) - c_{\ell, m, f}) q_{\ell, m, f, t} \right] \quad (1)$$

- $Q_{f, t} = (Q_{\ell, f, t})_{\ell \in \mathcal{L}}$ is a vector of firm capacities of different types ℓ at period t .
- $q_{\ell, m, f, t}$ are realizations of demand at period t . Capacity is shared within demand of types ℓ so that realized demand obeys $\sum_{m \in \mathcal{M}} q_{\ell, m, f, t} \leq Q_{\ell, f, t}$. The realized demand vector is drawn from a distribution $g(\cdot | p_i, Q_i, h)$
- h_t is a state variable affecting demand and evolving stochastically. We assume there is an absorbing state \bar{h} in which there is zero demand, effectively serving as a terminal condition for the pricing problem.
- $p_i(Q_i, h) = (p_{\ell, m, f}(Q_i, h))_{\ell \in \mathcal{L}, m \in \mathcal{M}}$ is the vector of firm prices contingent on the current state.

Let $J(p_i)$ denote the objective function of the problem. A brute force approach to solving this problem via gradient ascent is to simulate the expectation to a high-degree of precision and take the gradient of this expectation with respect to p_i at each iteration of the algorithm. Because the expectation is over many long demand *paths*, the variance is high, requiring many draws to achieve high precision and hence an unfeasibly high amount of computation time.

Fortunately, the problem fits neatly into the framework of policy gradient algorithms with function approximation, as proposed by Sutton, McAllester, Singh, and Mansour (1999). A key feature of this approach is the use of *stochastic gradient ascent*, which can use a noisy approximation of the gradient at each iteration, so long as this gradient takes the form of an expectation. This noisy approximation can use significantly less simulation draws, and hence, computational resources, than a precise gradient would. Using a simplified notation, it is

straightforward to derive the gradient in expectation form. In the following, let $\tau = (q_t)_{t=0}^{\infty}$ be a **trajectory** of realized vehicle demands, with associated distribution $\tilde{g}(\tau|p)$:

$$\begin{aligned}
\nabla_p J(p) &= \nabla_p \mathbb{E}_{\tau} \left[\sum_{t=0}^{\infty} \beta^t (p(Q_t) - c)' q_t \right] \\
&= \nabla_p \sum_{\tau} \tilde{g}(\tau|p) \left[\sum_{t=0}^{\infty} \beta^t (p(Q_t) - c)' q_t \right] \\
&= \sum_{\tau} \tilde{g}(\tau|p) \left[\sum_{t=0}^{\infty} \beta^t q_t \right] + \sum_{\tau} \nabla_p \tilde{g}(\tau|p) \left[\sum_{t=0}^{\infty} \beta^t (p(Q_t) - c)' q_t \right] \\
&= \sum_{\tau} \tilde{g}(\tau|p) \left[\sum_{t=0}^{\infty} \beta^t q_t \right] + \sum_{\tau} \tilde{g}(\tau|p) \nabla_p \log \tilde{g}(\tau|p) \left[\sum_{t=0}^{\infty} \beta^t (p(Q_t) - c)' q_t \right] \\
&= \sum_{\tau} \tilde{g}(\tau|p) \left[\sum_{t=0}^{\infty} \beta^t q_t + \nabla_p \log \tilde{g}(\tau|p) \left(\sum_{t=0}^{\infty} \beta^t (p(Q_t) - c)' q_t \right) \right] \\
&= \mathbb{E}_{\tau} \left[\sum_{t=0}^{\infty} \beta^t q_t + \nabla_p \log \tilde{g}(\tau|p) \left(\sum_{t=0}^{\infty} \beta^t (p(Q_t) - c)' q_t \right) \right]
\end{aligned}$$

Two steps in the above derivation merit additional explanation. Firstly, the expectation is converted into a sum weighted by the probability of the demand realizations $\tilde{g}(\tau|p)$. Second, the key to making the gradient into an expectation is the identity $\nabla_p \tilde{g}(\tau|p) = \tilde{g}(\tau|p) \nabla_p \log \tilde{g}(\tau|p)$.

We can further simplify this expression by decomposing the probability of a trajectory into a product of conditional probabilities:

$$\begin{aligned}
\nabla_p \log \tilde{g}(\tau|p) &= \nabla_p \log \left(\overbrace{g(q_0|p, Q_0)}^{\text{Conditional demand}} \overbrace{t(Q_1|q_0, Q_0)}^{\text{Capacity transition}} g(q_1|p, q_1) t(Q_2|q_1, Q_1) \dots \right) \\
&= \nabla_p \log \left(\prod_{t=0}^{\infty} g(q_t|p, Q_t) t(Q_{t+1}|q_t, Q_t) \right) \\
&= \sum_{t=0} \nabla_p \log g(q_t|p, Q_t) + \underbrace{\nabla_p \log t(Q_{t+1}|q_t, Q_t)}_{=0} \\
&= \sum_{t=0} \nabla_p \log g(q_t|p, Q_t)
\end{aligned}$$

When approximating this expectation via simulation, it is not necessary to simulate a trajectory of infinite length, due to the absorbing state \bar{h} (see Section 4). Let s index each simulation, with S total simulation draws and denote T_s as the final period of each simula-

tion, when $h_{T_s} = \bar{h}$. We then obtain the simulated approximation to the expected gradient:

$$\tilde{\nabla}_p J(p) = \frac{1}{S} \sum_{s=1}^S \left[\sum_{t=0}^{T_s} (\beta^t q_{s,t} + \nabla_p \log g(q_{s,t}|p, Q_t)) \sum_{t=0}^{T_s} ((p - c)' q_{s,t}) \right] \quad (2)$$

The next issue is that $p_i(Q_i, h)$ is a *function*, which is equal in size to the Cartesian product of the state-space and the price space (both of which are high-dimensional), and thus is not computationally feasible to optimize on directly. We can thus use a simplified parameteric pricing function $p_{\theta_i}(Q_i, h)$ over which the optimization will be performed, as in the second equilibrium of Section 4.3. In that case, the gradient takes the form:

$$\tilde{\nabla}_{\theta} J(\theta) = \frac{1}{S} \sum_{s=1}^S \left[\sum_{t=0}^{T_s} (\beta^t (\nabla_{\theta} p_{\theta}) q_{s,t} + (\nabla_{\theta} p_{\theta}) \nabla_p \log g(q_{s,t}|p, Q)) \sum_{t=0}^{T_s} ((p_{\theta} - c)' q_{s,t}) \right] \quad (3)$$

where, the gradient w.r.t. p is transformed into a gradient w.r.t. θ via the chain rule.

Finally, the stochastic gradient ascent algorithm works by repeatedly updating the parameters as follows:

$$\theta_i^{k+1} = \theta_i^k + \alpha_k \tilde{\nabla}_{\theta} J(\theta) \quad (4)$$

There are additional performance improvements and implementation details that can be found in Section A.3 of the online appendix. To better understand the behavior of the algorithm, the next section demonstrates the effectiveness and behavior of the algorithm in a simple linear differentiated Bertrand setting, where the algorithm's results can be compared to the analytical solution.

2.1 Simplified benchmark application

To better understand the properties of the reinforcement learning algorithm, this section demonstrates how the algorithm performs in a highly simplified setting of two firms competing under linear differentiated Bertrand demand without capacity constraints. While this algorithm would never be used in such a simple setting in practice, it provides a convenient benchmark due to its easily computed analytical solution.

The demand for each firm $i \in \{1, 2\}$, $j \neq i$ is:

$$q_i(p_i, p_j) = A_i - a_i p_i + b_i p_j$$

Since the algorithm relies on demand being stochastic, we will assume that the demand each firm faces each period is scaled by the total market size M_t , which is drawn each period from a Poisson distribution with rate λ . We thus have:

$$q_{f,t}(p_i, p_j) = M_t q_i(p_i, p_j) \sim \text{Poisson}(\lambda q_i(p_i, p_j))$$

Each firm solves the following maximization problem:

$$\max_{p_i} \mathbb{E}[(p_i - c_i)q_{i,t}(p_i, p_j)] \Rightarrow \max_{p_i} (p_i - c_i)\lambda q_i(p_i, p_j)$$

The simplified maximization problem does not depend on the stochastic nature of demand, and can thus be reduced to the well-known Nash-Bertrand equilibrium, characterized by the following first-order conditions:

$$\begin{aligned} A_i - 2a_i p_i + b_i p_j + a_i c_i &= 0 \\ 2p_i^* &= \frac{1}{a_i}(A_i + b_i p_j) + c_i \end{aligned}$$

The equilibrium is the solution to the resulting system of linear equations, providing an easily computable baseline with which to compare the results obtained from the reinforcement learning algorithm. Next, we can derive the stochastic gradient for this simplified model.

We return to the firm's problem, this time decomposing the expectation into a sum:

$$\begin{aligned} \max_{p_i} \mathbb{E}[(p_i - c_i)q_{i,t}(p_i, p_{j,k})] \\ \max_{p_i} \sum_{q_i \in \mathbb{N}} g(q_i | p_i, p_{j,k}) [(p_i - c_i)q_i] \end{aligned}$$

Where $g(q_i | p_i, p_{j,k})$ is the probability mass function of a Poisson distribution with rate $\lambda q_i(p_i, p_j)$. Importantly, the derivative of this PMF with respect to the firm prices is analytically tractable.

As before, we can define the stochastic gradient:

$$\begin{aligned}
\nabla_{p_i} J(p_i) &= \nabla_p \sum_{q_i \in \mathbb{N}} g(q_i | p_i, p_{j,k}) [(p_i - c_i) q_i] \\
&= \sum_{q_i \in \mathbb{N}} g(q_i | p_i, p_{j,k}) q_i + \sum_{q_i \in \mathbb{N}} \nabla_p g(q_i | p_i, p_{j,k}) [(p_i - c_i) q_i] \\
&= \sum_{q_i \in \mathbb{N}} g(q_i | p_i, p_{j,k}) q_i + \sum_{q_i \in \mathbb{N}} g(q_i | p_i, p_{j,k}) \nabla_p \log g(q_i | p_i, p_{j,k}) [(p_i - c_i) q_i] \\
&= \sum_{q_i \in \mathbb{N}} g(q_i | p_i, p_{j,k}) \left[q_i + \nabla_p \log g(q_i | p_i, p_{j,k}) [(p_i - c_i) q_i] \right]
\end{aligned}$$

The stochastic gradient then involves taking S draws of q given the current $p_{i,k}$ and $p_{j,k}$ and forming:

$$\tilde{\nabla}_{p_{i,k}} J(p_{i,k}) = \frac{1}{S} \sum_{s=1}^S \left[q_i^s + \nabla_{p_{i,k}} \log g(q_i^s | p_{i,k}, p_{j,k}) [(p_{i,k} - c_i) q_i^s] \right]$$

The iteration then proceeds as follows:

$$p_{i,k+1} = p_{i,k} + \alpha_k \tilde{\nabla}_{p_{i,k}} J(p_{i,k})$$

To demonstrate the algorithm, we use the baseline parameters shown in Table 1. These parameters result in an equilibrium with $p_1^* = 4.8637$ and $p_2^* = 3.9384$.

A_1	A_2	a_1	a_2	b_1	b_2	c_1	c_2	λ
4	5	0.7	0.9	0.5	0.3	1.2	0.7	2

Table 1: Differentiated Bertrand demand parameters

In this simple setting, we will show how the algorithm converges to a solution, and how convergence is affected by various training parameters. First, we note that the algorithm allows the update size α_k to vary with each iteration k . In general, it is beneficial to start with a relatively high α_k and let it decay as k increases, which efficiently trades off the speed and accuracy of the algorithm. However, care should be taken not to decrease α_k too fast, lest the algorithm become stuck (Sutton and Barto, 2018). A common approach is exponential decay, which sets:

$$\alpha_k = \alpha e^{-\delta k}$$

with the training parameter δ governing the rate of decay.

For the baseline training parameters, we use a baseline learning rate of $\alpha = 0.02$, a decay rate $\delta = 0.02$, total simulation draws $S = 5$ and a total of $K = 200$ training iterations.

For each set of training parameters we run the algorithm once. The error of the algorithm’s learned prices (denoted p_1^K and p_2^K) relative to the true equilibrium of $p_1^* = 4.8637, p_2^* = 3.9384$ are plotted in Figure 1 for a single run of the algorithm under various training parameters. For Table 2, the algorithm is run 100 times for each set of training parameters, and the mean and standard deviation of the final training error are compared.

Parameter set	$\mu(p_1^K - p_1^*)$	$\mu(p_2^K - p_2^*)$	$\sigma(p_1^K - p_1^*)$	$\sigma(p_2^K - p_2^*)$
Baseline	-0.0064	-0.0012	0.0274	0.0291
More simulation draws, $S = 100$	-0.0017	-0.0011	0.0073	0.0055
Lower training rate, $\alpha = 0.01$	-0.0857	-0.0448	0.0298	0.0264
Faster decay rate, $\delta = 0.03$	-0.0159	-0.0129	0.0346	0.0289
High initial price, $p_{\text{init}} = 6$	0.0030	0.0065	0.0293	0.0268
Higher Poisson rate $\lambda = 6$	-2.7739e-03	2.1896e-05	0.0350	0.0324

Table 2: Performance of various training parameter sets over 100 runs

With the baseline parameters, the stochastic nature of the algorithm is evident, as the prices converge neither smoothly nor monotonically towards the true equilibrium. Both increasing the simulation draws and reducing the training rate have the effect of lowering the variance of the algorithm, though the latter comes at the expense of slower training in terms of iterations. The increased simulation draws also result in a lower variance in the final training error.¹¹

A faster decay rate also appears to stabilize the training faster, but results in a much worse performance on average, suggesting the decay is too fast and causes the algorithm to get stuck before converging.

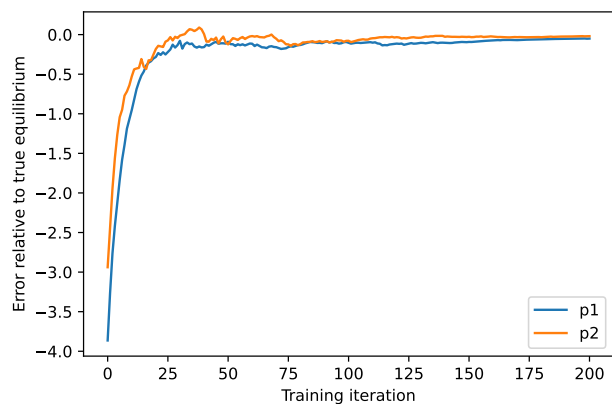
When the initial price is set above the true equilibrium rather than below, the error tends to be positive rather than negative, suggesting the algorithm retains some path-dependence from the initial guess.

In addition to changing the training parameter, we can also check what happens when we alter the model parameter λ , the Poisson arrival rate of consumers. As shown before, this has no effect on the true equilibrium prices, but does affect the performance of the algorithm, increasing both the rate of learning and the variance of the gradient. In practice, this may mean that the optimal training parameters are dependent on the properties of the demand system.

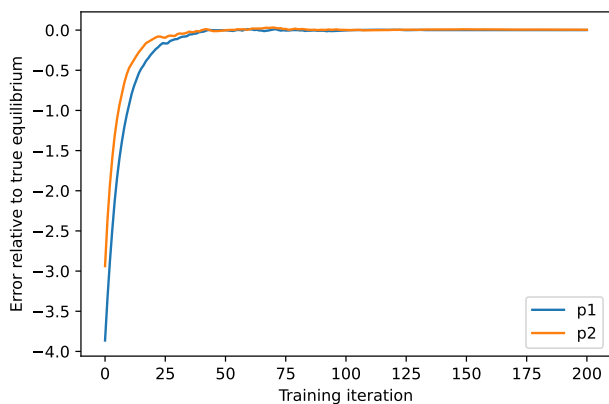
Finally, given that the algorithm trains on market outcomes, it is possible that changes in

¹¹As described in Section A.3 of the online appendix, we can reduce the variance of the gradient by subtracting the expected profit from the realized profit weight. The modified gradient is:

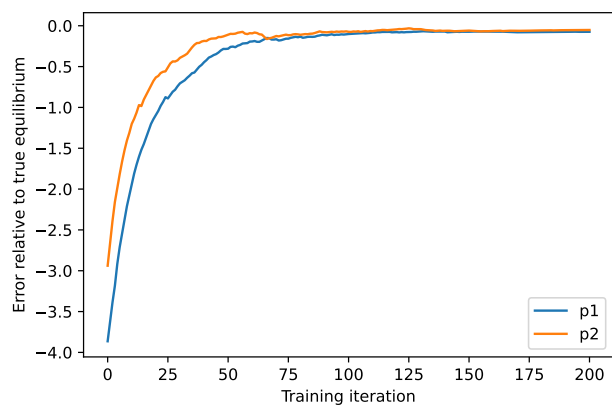
$$\tilde{\nabla}_{p_{i,k}} J(p_{i,k}) = \frac{1}{S} \sum_{s=1}^S \left[q_i^s + \nabla_{p_{i,k}} \log g(q_i^s | p_{i,k}, p_{j,k}) [((p_{i,k} - c_i) q_i^s) - (p_{i,k} - c_i) \tilde{\lambda}(p_{i,k}, p_{j,k})] \right]$$



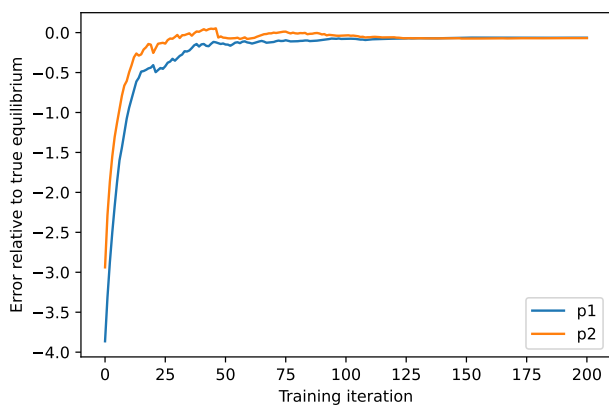
(a) Baseline



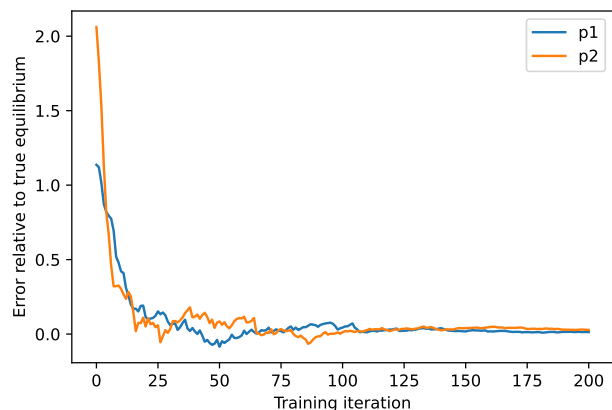
(b) More simulation draws, $S = 100$



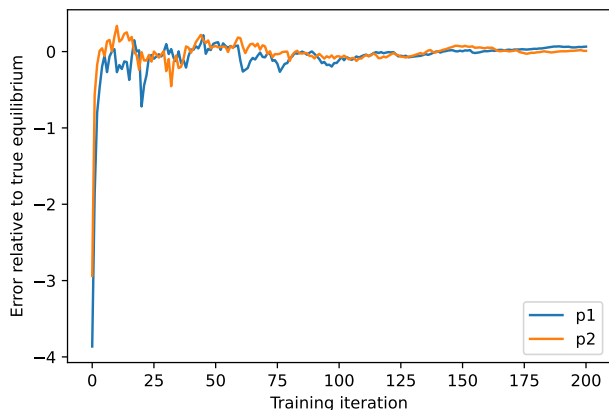
(c) Lower training rate, $\alpha = 0.01$



(d) Faster decay rate, $\delta = 0.03$



(e) High initial price, $p_{init} = 6$



(f) Higher Poisson rate $\lambda = 6$

Figure 1: Effect of training parameters, single training run

the competitiveness of each firm may affect the performance of the algorithm, since a less competitive firm sees less demand and may thus learn more slowly, and a higher demand firm may learn faster but with higher variance, both in accordance with the pattern shown in subfigure 1f and the scenario with $\lambda = 6$ in Table 2.

In Figure 2, the algorithm is run on a 21 by 21 grid of different costs for each firm, with 100 runs of the algorithm for each cell in the grid. The training parameters used for this figure are $K = 200$, $\alpha = 0.02$, $\delta = 0.02$, $S = 100$ and $p_{init} = 1$. The range of costs considered is chosen such that equilibrium firm quantities for one of the two firms approaches 0 when their cost is maximal and the rival's cost is minimal, as can be seen in subfigures 2a and 2b.

Looking at the mean absolute error, we see that when costs are such that a firm's equilibrium demand is low, the final error is larger on average. This suggests a slower rate of learning which fails to converge in the $K = 200$ iterations. On the other hand, the variance of the training error is increasing in the firm's equilibrium demand, which is consistent with faster but higher variance learning.

The main takeaways from this simplified exercise are that:

1. Given appropriate parameters, convergence appears robust, although some level of error in the final result is likely.
2. The optimal set of training parameters is highly specific to the characteristics of the demand faced by each firm.
3. Certain parameters are more harmful than others when chosen poorly. This is particularly true for the decay parameter.

Because of these observations, the algorithm used in the counterfactual uses conservative parameters, favoring (i) a slower learning rate, (ii) a higher count of simulation draws, leveraging the power of parallel computing, and (iii) foregoing decay in the learning rate altogether, to avoid any risk of the algorithm getting stuck.

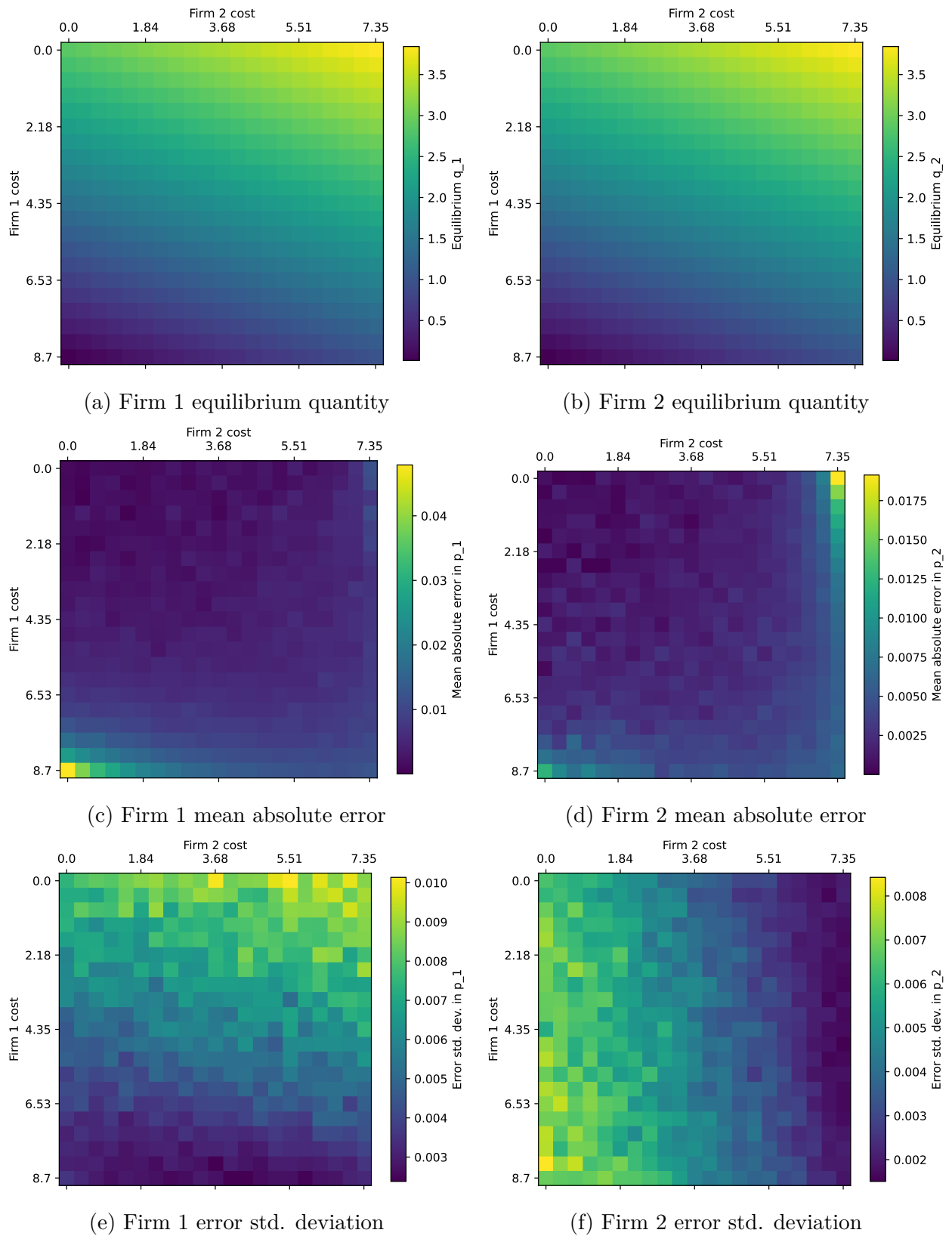


Figure 2: Algorithm performance over grid of costs, 100 runs per cell

Figure 3: Errors are the difference between true equilibrium prices and the price obtained by the algorithm after $K = 200$ iterations for each run.

3 Empirical setting and Data

3.1 Micromobility in Washington, D.C.

Dockless vehicle sharing platforms rent out fleets of personal mobility vehicles (usually electrically powered kick scooters or bicycles), variously referred to as "micro-mobility" or "dockless vehicles", to contrast with station-based bikesharing systems. Users on these platforms can view available vehicles on a mobile application and unlock them (generally for a fixed fee), after which they are charged a per-minute fee for their use. This business model was popularized by the rival platforms Bird and Lime, both of which were founded in 2017. The industry saw incredibly rapid growth, with 96 million trips nationwide in 2019. This already eclipsed the more established station-based bikeshare systems, which saw only 40 million trips in the same year.¹² The industry also saw numerous firms enter the market, including the two major rideshare platforms Lyft and Uber. The latter operated the Jump brand of scooters and bikes during the time period under study.

In Washington, D.C., the District Department of Transportation (DDOT) regulates this market by issuing permits to operate in the market. The price controls I study are specified in the terms and conditions of operation, which stipulate that firms "shall not charge customers at a more expensive rate for rental of dockless vehicles, or impose an additional fee, regardless of the trip origin and destination" (DDOT, 2019, DDOT, 2020) which effectively restricts firms from deploying more sophisticated price menus which vary directly with trip origin and destination. The same provision makes exceptions for "incentives for active rebalancing of dockless vehicles." However, there is no clear definition of what would constitute incentives for rebalancing as opposed to a more flexible pricing system, which may also be extracting greater rents. I thus take the requirement of price uniformity to be total in the status quo, and relax this requirement in the counterfactual exercise.

The operating permits issued by DDOT also set restrictions on fleet sizes and requirements on scooter deployment. In 2019, DDOT issued permits to a total of eight firms, with fleet size limits under one thousand per operator and vehicle type. It is likely that relaxing capacity restrictions would reduce the importance of pricing efficiency, since supply will be generally more abundant. However, Erkan, Stamatopoulos, Agarwal, and Muthuraman (2022) provide empirical and theoretical evidence that firms over-provision capacity, which may justify such capacity restrictions. There may also be negative externalities from vehicle deployment, due

¹²<https://nacto.org/wp-content/uploads/2020/08/2020bikesharesnapshot.pdf>

to the sidewalk space they take up, but the cost of this externality has not been quantified. I thus take the capacity restrictions as given in this paper.

The regulations also require minimum deployments of vehicles in each of the city’s eight wards in the morning of every day. The 2019 regulations require a minimum of six vehicles in each ward (DDOT, 2019), while the 2020 regulations (DDOT, 2020) require a minimum of 20. In addition, the 2020 regulations require a minimum of 400 vehicles in prescribed “Equity Emphasis Areas”, as well as no more than 1,000 vehicles in the Central Business District. However, the small fleet sizes (see Table 3) at the time the regulation came into force appear incompatible with these regulations, so they were unlikely to have been enforced. Non-enforcement was evident during the onset of the COVID-19 pandemic, when the regulator planned to downsize the number of permitted firms from eight down to four. From the author’s own continued data collection, all firms were able to continue operating throughout the year.¹³

3.2 Data

As part of the conditions of operating in Washington D.C., firms are required to publish an API disclosing real-time GPS information on each of their available vehicles. I collected this data at minute-level frequency and use the time period from November 26th, 2019, and March 8th, 2020.¹⁴ Certain firms included persistent vehicle identifiers in their APIs, which was used to reconstruct vehicle trips by observing the disappearance of a vehicle in one location and its reappearance in another location at a later time. These vehicle trips were then filtered to remove trips that seemed too short, too long, or whose implied speed exceeded the capabilities of the vehicles. The average number, distance, and price of the trips for the relevant firms can be found in Table 3.

To complement the micromobility data, I also use the Transportation Planning Board (TPB) Regional Travel Survey of 2017/2018. Respondents were asked to log all their trips in the 24 hours preceding their response. The raw trip data includes the origin and destination, as well as the start and end times. The TPB reports origins and destinations in terms of Traffic Analysis Zones (TAZ), which aggregate census blocks. For tractability of the estimation and counterfactual computation, I further aggregate TAZs into a grid of 22 cells covering

¹³See: <https://www.washingtonpost.com/transportation/2020/02/27/lime-bird-bolt-razor-lose-appeals-operate-scooters-dc/>

¹⁴The data collection was continued for most of 2020 and 2021, but the shorter time period was chosen to avoid the impact of the COVID-19 pandemic.

Variable	Bird	Jump_Bike	Jump	Lime	Lyft	Razor	Skip	Spin
Daily trips, mean	-	-	987.31	-	-	15.94	-	469.81
Trip distance, mean	-	-	1349.71	-	-	1676.24	-	1387.75
Trip price, mean	-	-	3.18	-	-	4.95	-	3.9
Total vehicle stock	230	392	338	471	410	284	624	447
Stock of vehicles, Q(0.1)	2	5	3	6	7	5	7	6
Stock of vehicles, median	7	15	12	16	16	11	24	18
Stock of vehicles, Q(0.9)	23	34	34	41	33	23	56	39
Stock difference , Q(0.1)	0	0	0	0	0	0	0	0
Stock difference , median	0	1	1	1	1	0	1	0
Stock difference , Q(0.9)	2	3	2	3	2	0	3	2
Walking distance, Q(0.1)	190.9	170.25	181.41	153.02	161.83	184.08	139.17	165.42
Walking distance, median	389.42	264.38	308.31	258.39	255.21	292.11	254.6	255.25
Walking distance, Q(0.9)	647.78	434.66	543.68	448.21	407.57	430.54	455.52	422.98

Q(n) denotes the n-th quantile. Walking distance in meters. Stock of vehicles is within a grid cell. Stock differences are in absolute terms.

Table 3: Descriptive Statistics

the areas seeing most of the micromobility traffic in Washington, D.C. The resulting grid is shown in Figure 4.

I also discretize time into 10-minute intervals. This is greater than roughly two-thirds of trip durations, as seen in Figure A.1 in the online appendix. In the model, it will be assumed that trips only last for a single one of these intervals, thereby avoiding the additional complications to the state space of modeling vehicle transit times.

As the trip data can only be reconstructed for a subset of the firms in the market, I complement it with a time series of the stock of vehicles available in each location, as well as the difference in stocks at the 10-minute frequency (hereinafter referred to as **stock difference** data). As explained later in Section 5.1.3, the stock difference from one period to the next is informative about trips being taken, and will thus be used in estimating demand-side parameters.

An additional input to the demand model is the expected walking distance from a consumer to the nearest vehicle of each firm. To compute this for each time period, I compute the distance to the nearest vehicle over a finer grid of points (as can be seen in Figure A.2 in the online appendix), spaced 100 meters apart. I then average this walking distance over each grid cell. The tenth percentile, median, and ninetieth percentile of the stock, (absolute value of) stock difference, and walking distance variables are displayed in Table 3.

Finally, to allow the demand model to capture differences in transportation patterns over the course of the day, I subset the day into four distinct times of day, over which most demand parameters will vary. As seen in Figure 5, most trips take place between 6 a.m. and 9 p.m. I thus take four intervals of 3.5 hours each, demarcated by the lines in the figure. This

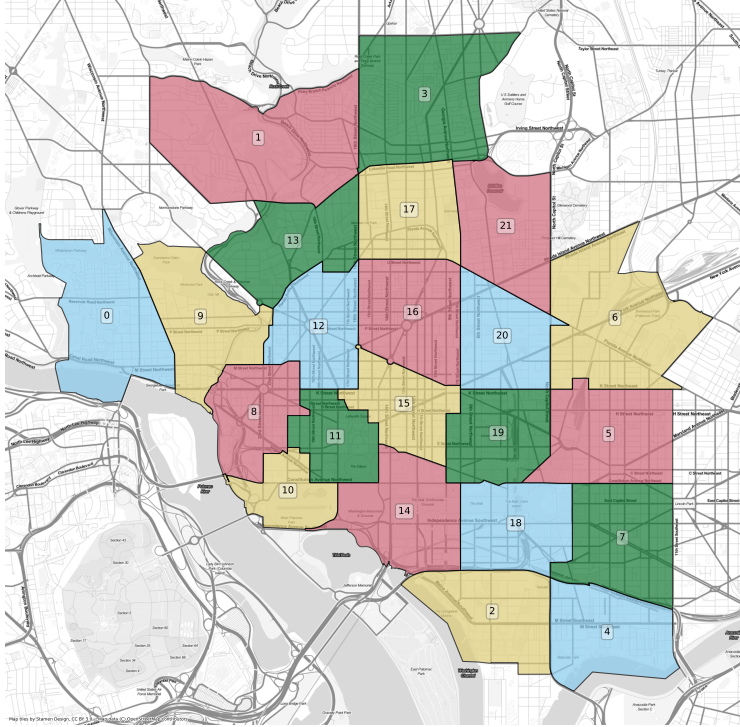


Figure 4: Location grid

Note: coloration is added solely for visual contrast.

roughly captures morning, noon, afternoon, and evening patterns of activity. Furthermore, I limit the data to weekdays only, as weekends have significantly different traffic patterns. The descriptive statistics in Table 3 reflect these restrictions, with the final panel amounting to 4,579 periods worth of observations across 8 products (7 firms) and 22 locations.

4 Model

The model consists of a demand process, which specifies arrival rates of consumers for each trip as a function of the prices and the current distribution of vehicles and a pricing equilibrium concept between the firms. As explained in the previous section, the model is set in discrete time with 10-minute intervals indexed by t . I assume that vehicles used for a trip in period t are available for use again at the destination in period $t + 1$. Space is discretized into a set of locations $\ell \in \mathcal{L}$. A trip consists of an origin ℓ and destination m . The firms are denoted $f \in \mathcal{F}$.

In addition, each day is divided into *times of day* h , which capture the major fluctuations in demand patterns over the course of the day (as described in Section 3.2 above). To simplify

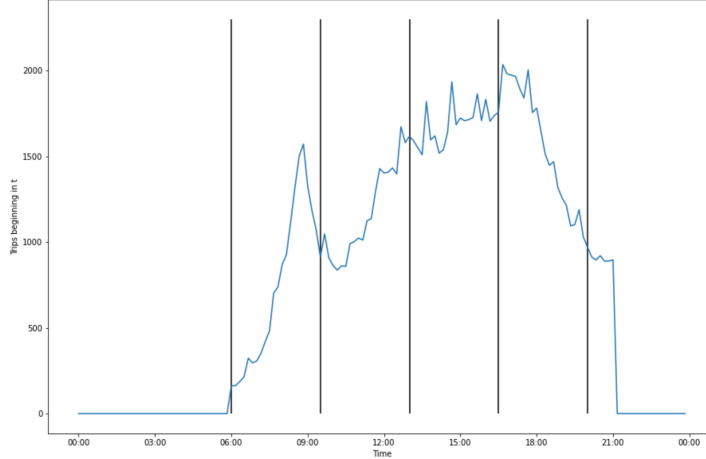


Figure 5: Density of trip times and discretization of time

dynamics, I assume the times of day evolve stochastically via a Markov process. Concretely, the hours of day each last 3 and a half hours, or 210 minutes, equivalent to 21 periods t . If in time t , the current time of day is $h_t = 1$, the morning, there will be a $\frac{1}{21}$ chance that the time of day in the next period $t + 1$ will be $h_{t+1} = 2$, the noon time of day, and a $\frac{20}{21}$ chance that the time of day remains the same at $h_{t+1} = 1$. This formulation leads to mean durations of each time of day of 210 minutes, but avoids needing to condition the firm’s problem on the more granular time index t . If the current time period is $h_t = 4$, the evening, then the time of day can transition to $h_{t+1} = \bar{h}$, the end of day, which is an absorbing state with no demand activity.

4.1 Demand

At the beginning of each t , each firm’s prices for each trip $p_{\ell,m,f,t}$ and the expected walking distance to the nearest vehicle in each location $\mathbb{E}(w_{\ell,f,t})$ are observed by consumers.

Then, within period t , the number of consumers desiring a trip (ℓ, m) is given by a Poisson arrival process $\lambda_{\ell,m,h}$, which depends on the time of day h . Consumers then choose among $\mathcal{F} \cup \{0\}$, where 0 represents the outside option (either an alternative mode of transportation or no trip at all). It is assumed that consumers exit the market when choosing the outside option and will not wait for better opportunities to take a scooter. In case of stockouts, consumers are served on a first-come first-serve basis.¹⁵ The utility of each of the inside

¹⁵Empirically, I observe very few (less than 1% of observations) possibilities of true stockouts based on the number of trips taken and the number of remaining vehicles, so this possibility is ignored in the demand estimation. However, since this fraction depends on firm prices, the possibility of stockouts is explicitly

choices is given by:

$$u_{\ell,m,t,f} = \underbrace{\gamma_{\ell,m,h} + \xi_f - \alpha p_{\ell,m,f} - \chi \mathbb{E}(w_{\ell,f,t})}_{\delta_{\ell,m,t,f}} + \eta_{in} + \sigma \epsilon_{t,f}, \quad (5)$$

where $\gamma_{\ell,m,h}$ captures the utility of the trip from ℓ to m , ξ_f is a firm-specific utility, $p_{\ell,m,f}$ is the price of the trip with firm f , $\mathbb{E}(w_{\ell,f,t})$ is the expected walking distance to the nearest vehicle, η_{in} is a nesting shock that is common to all the inside choices, and $\epsilon_{t,f}$ is the idiosyncratic Type-1 Extreme Value shock. The combined mean utility $\delta_{\ell,m,t,f}$ is common to all consumers, while the shocks are drawn independently across consumers. Note that the trip utility parameters $\gamma_{\ell,m,h}$ are indexed by time of day in order to capture changes in the intensity and direction of traffic flows appropriately over the course of the day.

For tractability of the estimation, the $\gamma_{\ell,m,h}$ are constrained to a simpler specification:

$$\gamma_{\ell,m,h} = \gamma_{O,\ell,h} + \gamma_{D,m,h} + \gamma_{1,h} \Delta_{\ell,m} + \gamma_{2,h} \Delta_{\ell,m}^2$$

where $\gamma_{O,\ell,h}$ is the value of a trip with *origin* ℓ at time of day h , $\gamma_{D,m,h}$ is the value of a trip with *destination* m , and $\Delta_{\ell,m}$ is the distance between locations ℓ and m . This constrained specification is similar to that in Buchholz et al., 2020, wherein the value of a trip relates to the value of leaving one location for another, as well as the distance of the trip, since other modes of transport contained in the outside option will generally have a cost proportional to the distance. Here, I add an additional term for the square of the trip distance, as the vehicles studied are bikes and scooters, which may be more uncomfortable than other modes of transportation over long distances.

The specification yields standard nested logit choice probabilities. The choice of firm, conditional on choosing the inside nest is:

$$P(\ell, m, t, f | inside) = \exp(\delta_{\ell,m,t,f}/\sigma) / \sum_{g \in \mathcal{F}} \exp(\delta_{\ell,m,t,g}/\sigma)$$

The probability of choosing the inside nest is:

$$P(inside) = \left(\sum_{g \in \mathcal{F}} \exp(\delta_{\ell,m,t,g}/\sigma) \right)^\sigma / \left(\left(\sum_{g \in \mathcal{F}} \exp(\delta_{\ell,m,t,g}/\sigma) \right)^\sigma + 1 \right)$$

Multiplying the arrival rates $\lambda_{\ell,m,h}$ by the choice probabilities yields demand arrival rates considered in the computation of counterfactual prices, in a similar fashion to Conlon and Mortimer (2013).

for each firm:

$$\tilde{\lambda}_{\ell,m,t,f} = \lambda_{\ell,m,h}P(\textit{inside})P(\ell, m, t, f|\textit{inside}) \quad (6)$$

4.2 Supply

I model each firm’s objective as seeking to maximize the discounted present value of its stream of revenues from each completed trip minus its operational costs of charging and redeploying vehicles. To do so it commits to both a relocation policy and a pricing policy, the latter of which can be contingent on both the time of day and possibly the stock of vehicles in each location.

I assume relocation occurs overnight jointly with the recharging of the vehicle batteries, and taking account of the firm’s own pricing policy.¹⁶ The pricing policy is set in the morning, taking into account the firm’s optimal relocation policy at the end of the day. As the relocation policy is held fixed in the counterfactuals, its formal exposition is relegated to Section A.2. I now focus on the firm’s pricing policy, which will be the subject of the counterfactuals.

The firm sets a pricing policy $p_f(Q_f, h)$ which specifies prices $p_{f,\ell,m}$ for each possible origin-destination pair (ℓ, m) as a function of the state variables Q_f , the firm’s own vector of vehicle stocks across all locations (thus emulating the surge pricing used by the popular ridehail platforms), and h , the time of day which indexes demand parameters. Formally, the firm pricing problem (omitting subscripts f) is:

$$\max_{p(Q,h)} \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{q,h|p} \left[\sum_{\ell \in \mathcal{L}} \sum_{m \in \mathcal{L}} (p_{\ell,m}(Q_t, h_t) - c_{\ell,m}) q_{\ell,m,t} + \mathbb{1}[h_t = \bar{h}](1 - \beta)V_R(Q_t) \right] \quad (7)$$

s.t.

$$\sum_{m \in \mathcal{L}} q_{\ell,m,t} \leq Q_{\ell,t}, \quad \forall t, \quad \forall \ell \in \mathcal{L} \quad (8)$$

$$Q_{\ell,t+1} = Q_{\ell,t} - \sum_{m \in \mathcal{L}} q_{\ell,m,t} + \sum_{k \in \mathcal{L}} q_{k,\ell,t} \quad (9)$$

$$Q_0 = \tilde{Q}_f^*$$

$$h_0 = 1$$

¹⁶The Jump data includes both vehicle identifiers and battery levels, allowing me to identify charging trips, which occur more frequently overnight, during which there is little to no demand.

The flow payoffs in the above problem simply sum the total profit over all possible origin-destination pairs, while the expectation operator is taken over the time of day states and the realized demand for each trip $q_{\ell,m,t}$. The latter is drawn from a Poisson process with arrival rate $\tilde{\lambda}_{\ell,m,t,f}(p_{\ell,m})$ defined by the demand system in Equation 6. The additional term $\mathbb{1}[h_t = \bar{h}](1 - \beta)V_R(Q_t)$ gives the firm a one-time end of day payoff of $V_R(Q_t)$ from the firm's solution to the relocation problem (defined in Equation A.2 in the online appendix).

Constraint 8 reflects the firm's capacity constraints, while constraint 9 is a law of motion capturing the evolution of the firm's vehicles stocks from the realized trips. The last two constraints specify initial conditions, where the initial stock Q_0 is assumed to be chosen by the firm optimally according to Equation A.1 in the online appendix. The evolution of the state variable Q_t is thus determined by the initial condition and the realizations of demand $q_{\ell,m,t}$.

4.3 Equilibrium

The equilibrium concept is Nash, where all firms simultaneously choose and commit to both their pricing and relocation policies. The assumption of commitment is reasonable in the context of transportation platforms, since policies are implemented in software and would be costly to alter in response to moment-to-moment changes in the state space.

I also assume that firms only observe the state of their own vehicle fleets, rather than that of competitors. While in principle, the open APIs used for this paper's data collection could also be used by firms to observe the state of each others' fleets, this is a feature unique to the Washington, D.C. market. In other markets, obtaining such data would require reverse-engineering the closed APIs of other firms, which presents technical and legal challenges. It is thus unlikely that firms would invest in pricing policies which condition on the state of rival fleets, only for use in a single market.

The formal equilibrium conditions are:

1. For each $f \in \mathcal{F}$, $p_f(Q, h)^*$ solves Equation 7, taking Q_f^* , p_{-f}^* , and Q_{-f}^* as given.
2. For each $f \in \mathcal{F}$, Q_f^* solves Equation A.1, taking p_f^* , p_{-f}^* , and Q_{-f}^* as given.

The first condition requires that firms solve their pricing problem by committing to a pricing schedule, taking their own relocation strategy and the pricing schedules and relocation

strategies of their rivals as given. The second condition requires that firms solve their relocation problem, taking their own pricing schedule and the pricing schedules and relocation strategies of their rivals as given.

We will also want to consider equilibria in which firms are constrained to a simpler pricing strategy. To model this, let firms be exogenously constrained to a parametric pricing strategy $p_{\theta_f}(Q_f, h)$, with $\theta_f \in \Theta_f$. This could represent a simple two-part tariff as in the status quo, or a very flexible function approximator such as a neural network which takes the state variables as inputs. In this case, the equilibrium conditions are slightly modified to include the restriction on the space of feasible policies:

1. For each $f \in \mathcal{F}$, $p_f(Q, h)^*$ solves 7, taking Q_f^* , p_{-f}^* , and Q_{-f}^* as given, **subject to** $p_f \in \{p_{\theta_f} : \theta_f \in \Theta_f\}$
2. For each $f \in \mathcal{F}$, Q_f^* solves A.1, taking p_f^* , p_{-f}^* , and Q_{-f}^* as given.

This will be referred to as the **parametric pricing equilibrium** in the following. This concept bears similarity to the Parametric Policy Iteration algorithm described by Rust, 2000 and implemented by Sweeting, 2013 and Bodéré, 2023, which involves solving a dynamic programming problem with a parametric approximation to the value function. Whereas the intent of that algorithm is to ease the researcher’s computational burden while providing a close approximation to the exact policy, the parametric pricing equilibrium described above allows for firms themselves to approximate the optimal strategy through a simplified pricing policy.

5 Estimation

5.1 Methodology

The estimation focuses on the demand model described in Section 4.1. The parameters of this model include the arrival rates of consumers $(\lambda_{\ell, m, h})$ and the utility parameters $(\gamma, \xi, \alpha, \chi, \sigma)$, referred to as θ_d for conciseness.

The arrival rates first are calibrated using survey data, and the consumer utility parameters are then estimated conditional on the calibrated arrival rates using a Generalized Method of Moments (GMM) estimator. Recall that the data consists of two broad types (see Table 3):

direct trip data for three firms and stock-difference data for all firms. The full estimation procedure is described in the following.

5.1.1 Arrival rates

The consumer arrival rates λ are essentially a measure of the non-deterministic market size. I interpret the market size as the total transportation demand across all modes between each origin-destination pair, differentiating by time of day.¹⁷ To calibrate this measure, I make use of the TPB travel survey described in Section 3.2. For each origin ℓ and destination m and time of day h , I group all the matching trips in the survey and sum the sampling weights. Recall that the time of day discretization results in four 3.5 hour periods corresponding to morning, midday, afternoon, and evening. To get arrival rates for the 10-minute periods indexed by t , I rescale the arrival rates down to obtain $\hat{\lambda}_{\ell,m,h}$. In addition, in order to avoid arrival rates of 0 that can arise due to some trips not being sampled in the survey, the arrival rates are all smoothed towards the mean, to. Formally:

$$\lambda_{\ell,m,h} = \left(\frac{9}{10} \hat{\lambda}_{\ell,m,h} + \frac{1}{10} \bar{\lambda}_h \right) \frac{1}{21}$$

where $\bar{\lambda}_h = \sum_{\ell \in \mathcal{L}} \sum_{m \in \mathcal{L}} \frac{\hat{\lambda}_{\ell,m,h}}{L^2}$.

5.1.2 Trip-level moments

The trip data consists of observations of $q_{\ell,m,f,t}$, which is the quantity of trips taken from ℓ to m with firm f 's vehicles at a time t . This is accompanied by $p_{\ell,m,f,t}$, the price of the trip, h_t the time of day, and $\mathbb{E}(w_{\ell,f,t})$, the expected walking distance to the nearest vehicle of the firm.

From the demand model in Section 4.1, I can compute the Poisson arrival rate $\tilde{\lambda}_{\ell,m,f,t}(p_{\ell,m,f,t}, h_t, \mathbb{E}(w_{\ell,f,t}); \theta_d)$ of $q_{\ell,m,f,t}$, conditional on the information above and the demand parameters θ_d . The Poisson arrival rate corresponds to the mean number of arrivals, so the first set of moments $\hat{g}_1(\theta_d)$ is constructed to match these as follows:

$$\hat{g}_{1,\ell,m,f}(\theta_d) = \frac{1}{T} \sum_{t=1}^T q_{\ell,m,f,t} - \tilde{\lambda}_{\ell,m,f,t}(\theta_d) \quad (10)$$

¹⁷This choice results in a very large share of the outside option. The nested logit specification is thus important to break the IIA property with respect to the outside share.

Let $\hat{g}_1(\theta_d)$ refer to the moments above stacked over ℓ , m , and f .

5.1.3 Stock-difference moments

To construct the stock difference moments, first recall the law of motion in 9 of a firm's stock levels, rearranged in terms of the stock difference variable:

$$\Delta Q_{\ell,f,t} = Q_{\ell,f,t+1} - Q_{\ell,f,t} = \underbrace{\sum_{k \in \mathcal{L}} q_{k,\ell,f,t}}_{\text{inflows}} - \underbrace{\sum_{m \in \mathcal{L}} q_{\ell,m,f,t}}_{\text{outflows}}$$

where the $q_{\ell,m,f,t}$ are in this case unobserved, but whose arrival rates are given by the respective $\tilde{\lambda}_{\ell,m,f,t}(\theta_d)$. The mean of the stock-difference variable is thus given by:

$$\mathbb{E}[\Delta Q_{\ell,f,t}] = \sum_{k \in \mathcal{L}} \tilde{\lambda}_{k,\ell,f,t}(\theta_d) - \sum_{m \in \mathcal{L}} \tilde{\lambda}_{\ell,m,f,t}(\theta_d) \equiv \mu_{\ell,f,t}(\theta_d)$$

The second set of moments $\hat{g}_2(\theta_d)$ is constructed to match the mean of the stock difference:

$$\hat{g}_{2,\ell,f}(\theta_d) = \frac{1}{T} \sum_{t=1}^T \Delta Q_{\ell,f,t} - \mu_{\ell,f,t}(\theta_d) \quad (11)$$

The crucial information in the stock difference data lies, however, in the covariances between the various stock differences. ¹⁸

Intuitively, the (negative) covariance in the stock evolution between two different locations will be caused by the traffic between the two locations. For example, suppose we have locations A , B , and C , and observe the stock of vehicles in each location. Suppose there are many vehicle trips between A and B , and between B and C , but very few between A and C , as illustrated in Figure 6. Next, we can examine the relationship in the stock differences across locations. A trip from A to B will cause a reduction in the stock in A and an increase in the stock in B , same as for any other pair of locations. Thus, the more trips between a pair of locations, the stronger the (negative) covariance in the stock fluctuations between

¹⁸Formally, the stock difference variable has a Skellam distribution, since it is effectively the difference between the two Poisson distributions which aggregate the inflows and outflows. Though this could be used to form a likelihood on a single stock difference time series, there is no known closed-form for the joint distribution of correlated Skellam processes, which is why the method of moments is preferred to a maximum likelihood estimator.

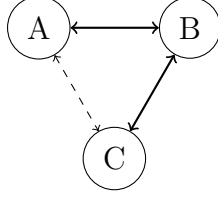


Figure 6: Illustration of covariances in stock differences

them will be. In this stylized example, we would expect to see a high covariance in the stock fluctuations between A and B and between B and C , but a low covariance in the stock fluctuations between A and C .

Formally, we can derive the covariance between the stock difference in two locations as:

$$Cov(\Delta Q_{\ell,f,t}, \Delta Q_{m,f,t}; \theta_d) = -(\tilde{\lambda}_{\ell,m,f,t}(\theta_d) + \tilde{\lambda}_{m,\ell,f,t}(\theta_d)), \quad \ell \neq m$$

The third set of moments $\hat{g}_3(\theta_d)$ is constructed to match the uncentered second moment of the stock difference *across different locations*:

$$\hat{g}_{3,\ell,m,f}(\theta_d) = \frac{1}{T} \sum_{t=1}^T \Delta Q_{\ell,f,t} \Delta Q_{m,f,t} - \left(Cov(\Delta Q_{\ell,f,t}, \Delta Q_{m,f,t}; \theta_d) + \mu_{\ell,f,t}(\theta_d) \mu_{m,f,t}(\theta_d) \right), \quad \ell \neq m \quad (12)$$

Lastly, recall that the trip data were cleaned to remove erroneous trips, which are likely caused by errors in the GPS data. The stock difference data is thus also likely to be polluted by the same GPS errors. To allow for this possibility, I assume that for each vehicle in a location ℓ , there is some arrival rate of a GPS error, which either adds 1 to or subtracts 1 from the stock difference variable ΔQ_{ℓ} . Thus, the variance of the stock difference variable *within a location* is inflated by the GPS error, proportionately to the level of the current stock:

$$Var(\Delta Q_{\ell,f,t}; \theta_d, \lambda_{GPS}) = \sum_{k \in \mathcal{L} \setminus \{\ell\}} \tilde{\lambda}_{k,\ell,f,t}(\theta_d) + \sum_{m \in \mathcal{L} \setminus \{\ell\}} \tilde{\lambda}_{\ell,m,f,t}(\theta_d) + 2\lambda_{GPS} Q_{\ell,f,t}$$

The fourth and final set of moments $\hat{g}_4(\theta_d, \lambda_{GPS})$ is constructed to match the uncentered second moment of the stock difference *within the same location*:

$$\hat{g}_4(\theta_d, \lambda_{GPS}) = \frac{1}{T} \sum_{t=1}^T \Delta Q_{\ell,f,t}^2 - \left(Var(\Delta Q_{\ell,f,t}; \theta_d, \lambda_{GPS}) + \mu_{\ell,f,t}(\theta_d)^2 \right) \quad (13)$$

5.1.4 Estimator details

The implementation of the GMM estimator follows Newey and McFadden (1994). The four sets of moments are stacked together¹⁹ as follows:

$$\hat{g}(\theta_d, \lambda_{GPS}) = \begin{pmatrix} \hat{g}_1(\theta_d) \\ \hat{g}_2(\theta_d) \\ \hat{g}_3(\theta_d) \\ \hat{g}_4(\theta_d, \lambda_{GPS}) \end{pmatrix}$$

The GMM estimator is then given by:

$$(\hat{\theta}_d, \hat{\lambda}_{GPS}) = \underset{\theta_d, \lambda_{GPS}}{\operatorname{argmin}} \hat{g}(\theta_d, \lambda_{GPS})' \hat{\Psi} \hat{g}(\theta_d, \lambda_{GPS}) \quad (14)$$

To increase efficiency, a two-step estimator is used, where in the first step $\hat{\Psi} = I$, and in the second step, the $\hat{\Psi}$ is chosen as the efficient weighting matrix, again according to Newey and McFadden (1994). The computation of standard errors follows the same reference.

5.1.5 Identification

The description of the moments above constructively describes how individual choice probabilities can be inferred from information on the stock of vehicles. This section further discusses how these choice probabilities identify the key parameters of the consumer utility function.

The price sensitivity parameter α is identified from the restriction of firm prices to a simple two-part tariff and from the assumption that the taste parameters for each firm ξ_f does not vary with trip distance. The fact that the majority of the firms all used scooters made by the same firm Ninebot, and that many initially used the same model, the Ninebot ES2, supports this assumption. Holding fixed the supply of each firm's vehicles, a firm with a lower price per mile will become *relatively* more attractive to consumers seeking longer rides. Formally, suppose we compare the choice probabilities of firms f and g across trips $A \equiv (\ell, m)$ and $B \equiv (\ell', m')$. Based on the formulation of utility in 5, we can write the log ratio of choice probabilities of each trip as:

¹⁹There are $3*(22)^2 = 1452$ moments in the first set, $8*22 = 176$ moments in the second set, $8*(21*22)/2 = 1848$ in the third set, and $8 * 22 = 176$ in the fourth set, for a total of 3,652 moments.

$$\begin{aligned}\log \frac{P(A, f)}{P(A, g)} &= u_{A,f} - u_{A,g} = \xi_f - \xi_g + \alpha(p_{A,f} - p_{A,g}) \\ \log \frac{P(B, f)}{P(B, g)} &= u_{B,f} - u_{B,g} = \xi_f - \xi_g + \alpha(p_{B,f} - p_{B,g})\end{aligned}$$

The difference in difference of choice probabilities is then directly given by the difference in difference of the prices:

$$\log \left(\frac{P(A, f)}{P(A, g)} / \frac{P(B, f)}{P(B, g)} \right) = \alpha((p_{A,f} - p_{B,f}) - (p_{A,g} - p_{B,g})) \quad (15)$$

Under the two-part tariff, this difference-in-difference of prices is a direct function of the differences in the price per mile and the differences in distances, which is naturally exogenous. The only remaining concern is whether the price per mile is correlated with the firm-specific utility ξ_f . Since firms make use of a two-part tariff, it is simpler for the firm to use the fixed part of the tariff to extract consumer utility from ξ_f , and that differences in the variable component can be attributed to differences in costs. We can thus identify α under this pricing regime.

The sensitivity to walking distance χ is identified from variation in the stock of vehicles. Implicitly this relies on the assumption that the idiosyncratic shocks $\epsilon_{t,f}$ are not autocorrelated across time, so that variation in the vehicle stock due to past demand can be taken as exogenous. The trip utilities γ act as a residual, and are identified from the overall size of the inside share for each combination of origin, destination, and time of day. This relies on the survey data (see Section 3.2) to accurately pin down the total potential market.

5.2 Results

In the following, I briefly present the parameter estimates of the demand system and compare them where possible to other studies.

Table 4 summarizes the key parameter estimates. As the trip origin and destination fixed effects are too numerous to list directly, the interquartile ranges are reported instead.²⁰

²⁰Though the standard errors for these parameters are not reported, it is worth noting that over 90% of parameters are precisely estimated with t-statistics over 2.

Transforming these into dollar equivalents, the interquartile range varies between \$3 and \$5, which is significant when considering that trip prices fall within a similar range.

Parameter		Morning	Midday	Afternoon	Evening
Trip distance (km):	$\hat{\gamma}_1$	0.2352 (0.1446)	0.2697 (0.1700)	0.5796 (0.0719)	0.3974 (0.0647)
Trip distance ² (km ²):	$\hat{\gamma}_2$	-0.0984 (0.0313)	-0.1620 (0.0402)	-0.1824 (0.0171)	-0.1222 (0.0121)
Trip origin IQR:	$\hat{\gamma}_O$	0.8553	1.1922	1.2435	1.1140
Trip destination IQR:	$\hat{\gamma}_D$	1.1975	0.9993	0.9082	1.5020
Price (\$):	$\hat{\alpha}$		-0.3034	(0.0155)	
Walking distance (m):	$\hat{\chi}$		-0.0149	(0.0005)	
GPS error rate:	$\hat{\lambda}_{GPS}$		0.0362	(0.0170)	
Nesting parameter:	$\hat{\sigma}$		0.4283	(0.0002)	

Table 4: Key parameter estimates

IQR: interquartile range

Note: Standard errors in parentheses.

The trip distance valuation is presented in Figure 7, which reports the trip value over different distances, holding all other factors fixed. In all cases, the utility is concave in distance and the estimates imply that trip distance quickly decreases utility after a few kilometers. It is also worth noting that longer trips are more desirable later in the day.

To assess the estimated trade-off between price and walking distance to the nearest vehicle, I compare the relative magnitudes of α and χ against two other recent studies on ride-sharing platforms, shown in Table 5. The first is Buchholz et al. (2020), who use data from an auction-based rideshare platform in Prague to estimate the trade-off between waiting time and price. The second is Rosaia (2020), who uses data from Uber and Lyft’s operations in New York City to estimate a demand model with parameters varying according to the ride length.

I transform my estimate of χ , the disutility of an additional 1 meter of expected walking distance, into a disutility of 1 minute of walking, which is approximately 85 meters at the average human walking speed. My estimates of the trade-off are in line with those of Rosaia (2020) for short trips, though both are far higher than in Buchholz et al. (2020). Note, however, that there are significant differences in local incomes of these studies which may explain some of the gap. The average income in Prague is approximately 21 thousand USD²¹, while the median incomes in Washington D.C., and New York City are approximately 88

²¹https://www.czso.cz/csu/czso/labour_and_earnings_ekon

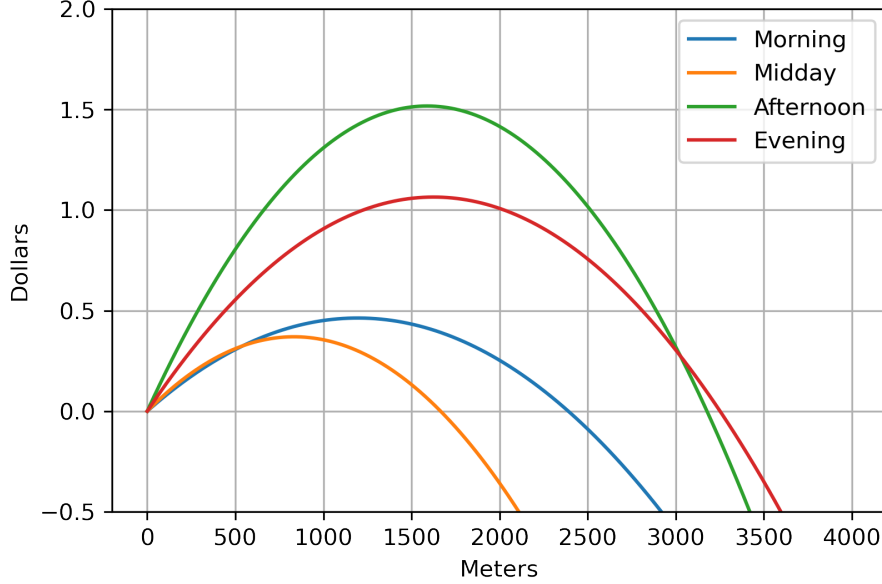


Figure 7: Value of trip distance

thousand USD²² and 68 thousand USD²³, respectively.

	Value
WTP for 1 minute walk	4.17 \$
Rosaia (2021) (≤ 15 min)	7.79 \$
Rosaia (2021) (30-35 min)	0.81 \$
Buchholz et al. (2020)	0.22 \$

Table 5: Comparison of time-price trade-off

Note: 1 minute walking time is equivalent to 85 meters

To further assess the estimated parameters, it is useful to examine the implied price elasticities. As there are a large number of effective products and prices, I present three different average elasticities with different weighting schemes in Table 6 and compare them with recent studies on rideshare platforms.

The unweighted average is a simple average across all possible trips, across all time periods in the data. The market size weighting uses the λ parameters calibrated from the travel survey. Finally, the trip weighting uses actual trips completed (for the 3 firms with trip data available). While this last measure yields an average elasticity below 1, which in principle is incompatible with firm profit maximization, this is in line with other studies on

²²<https://fred.stlouisfed.org/series/MEHOINUSNYA672N>

²³<https://fred.stlouisfed.org/series/MEHOINUSDCA672N>

transportation platforms, in particular Cohen et al. (2016), who find even lower estimated elasticities. This pattern may reflect the incentives of such firms to grow faster at the cost of short-run profit, in order to attract venture capital from investors. In order to deal with this fact, the counterfactuals also re-compute the optimal two-part tariffs under the objective of short-run profit maximization, to create a more level ground for comparison with the more sophisticated prices. Overall, the average elasticities appear reasonable and are contained within the range of other studies. It is worth noting that Buchholz et al. (2020) report bid-level elasticities on an auction platform with inter-driver competition which may account for the much larger elasticities.

	Elasticity
Unweighted	-2.51
Market size weighted	-2.05
Trip weighted	-0.74
Rosaia (2021) (≤ 15 min)	-0.55
Rosaia (2021) (30-35 min)	-1.08
Buchholz et al. (2020)	-4.37
Cohen et al. (2016), Overall	-0.55
Cohen et al. (2016), Los Angeles	-0.33

Table 6: Comparison of estimated average price elasticities

6 Counterfactuals

The objective of the counterfactuals is to evaluate the policy limiting pricing flexibility in Washington D.C. as described in Section 3.1. To this end, I re-compute the pricing equilibrium while allowing firms complete flexibility in setting prices for each origin-destination pair (ℓ, m) and each time of day h separately, a total of 1,936 prices for each of the 8 firms. Because of the high-dimensionality of the price schedule, I introduce a stochastic gradient descent algorithm based on policy gradient algorithms popular in the reinforcement learning literature.²⁴

6.1 Methodology

Before introducing the algorithm used to compute counterfactual prices, I begin with a few simplifying assumptions.

²⁴See Sutton and Barto (2018) for an overview.

Firstly, the relocation problem is not solved for in the counterfactual, due to the added computational burden it would impose. However, since firms could only do better by also resolving the relocation problem, the counterfactual results can be interpreted as a conservative lower bound on the efficiency gains. I thus take the status-quo relocation policy \check{Q}_f^* as fixed, setting it equal to the average distribution of vehicles in the morning time for each firm:

$$\check{Q}_f^* = \frac{1}{\sum_{t=1}^T \mathbb{1}[h_t = 1]} \sum_{t=1}^T Q_{f,t} \mathbb{1}[h_t = 1], \quad \forall f \in \mathcal{F} \quad (16)$$

Second, I ignore the costs of relocation by setting the relocation cost matrix C^r to 0. This is justified by the fact that vehicles must eventually be charged anyways, so that the relocation is given “for free” together with the charging. I then model the cost of charging as part of the trip marginal cost $c_{\ell,m}$, proportionately to the battery depletion. For the calibration, I assume a cost of recharge of \$4, which is roughly in line with the rates offered by Lime for charging vehicles, as seen in Figure A.3 of the online appendix. I then use a conservative range estimate of 10 kilometers, yielding a cost per-kilometer of 40 cents.²⁵

I now focus on the simplified firm pricing problem without the relocation component:

$$\max_{p_f(Q_f, h)} \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{q_f, h | p_f} \left[\mathbb{1}[h_t \neq \bar{h}] \sum_{\ell \in \mathcal{L}} \sum_{m \in \mathcal{L}} (p_{\ell, m, \theta_f}(Q_{f,t}, h_t) - c_{\ell, m, f}) q_{\ell, m, f, t} \right] \quad (17)$$

where the capacity constraint is implicitly accounted for in the distribution of the demand variable q . Let $J(p_f)$ denote the objective function of the problem. This problem is nearly identical to the general problem in Equation 1 in Section 2, with the only change being that $\mathcal{M} = \mathcal{L}$, that is, the types of demand are trips of different origins and destinations sharing a common space of locations, and capacity constraints are defined at the origin level.

Because of the cost calibration and the emerging nature of the market (with most firms being venture capital backed startups), the assumption of firm profit maximization may not be met under the prices charged by firms in the status quo. I thus re-compute the profit-maximizing equilibrium under the calibrated costs, and compare the equilibria where all firms symmetrically follow one of the three following pricing policies:

1. Two-part tariff $p_{\ell, m, f}(Q_f, h) = \theta_{f,1} + \theta_{f,2} \Delta_{\ell, m}$, where $\Delta_{\ell, m}$ is trip distance.

²⁵Tests of the Ninebot ES2, a popular electric scooter that was initially used by many of the firms in question, yielded range estimates of 16 kilometers. The conservative estimate allows for some deterioration of battery life.

2. Flexible, stock invariant prices $p_{\ell,m,f}(Q_f, h) = \theta_{f,\ell,m,h}$
3. Flexible, stock responsive prices $p_{\ell,m,f}(Q_f, h) = p_{\ell,m,h}(Q_f; \theta_f)$ represented by a neural network with parameters θ_f

The re-computed two-part tariffs are compared with the real two-part tariffs in Table A.2 in the online appendix. The second pricing policy allows prices to respond dynamically to changes in the stock of vehicles, similarly to the “surge pricing” algorithms implemented by many rideshare platforms. Note that the stock invariant pricing still allows prices to vary across time of day, so that they can still respond to the predictable changes in the patterns of demand over the course of the day.

6.2 Results

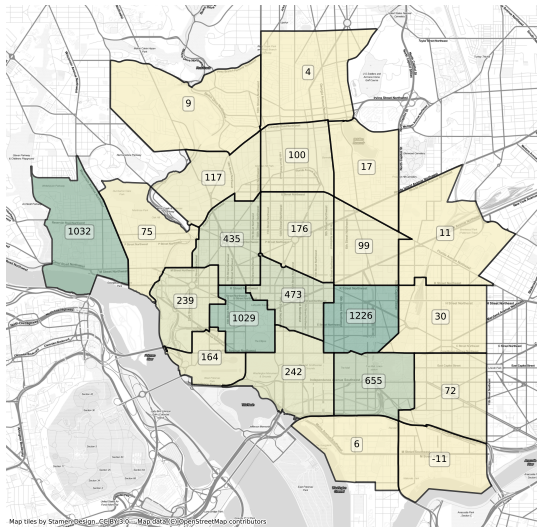
In the following, I focus on the change in profits, consumer surplus, and quantities (rides taken) across the three pricing policies. The overall outcomes at the daily level are given in Table 7.

	Two-Part	Flexible, stock invariant	Flexible, stock responsive
Profits	\$18343.94	\$24544.06	\$24257.04
	-	+33.80%	+32.23%
Quantities	9714.36	13675.66	13496.09
	-	+40.78%	+38.93%
Consumer Surplus	-	+\$14711.72	+\$14034.09

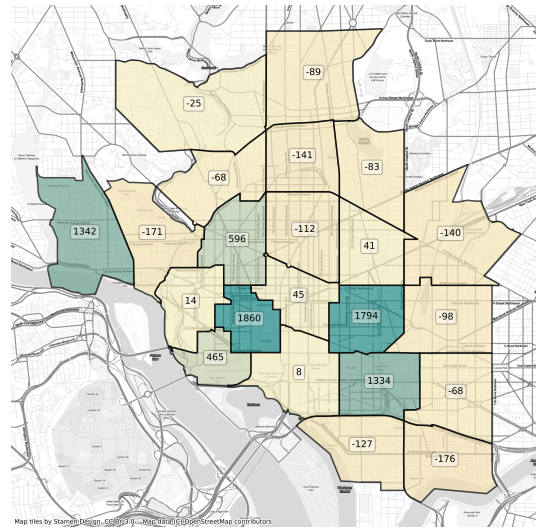
Table 7: Outcomes under alternative price equilibria, scaled to daily levels
 Note: since consumer surplus has a level normalization, the baseline is not reported.

Profits, quantities and consumer surplus are all substantially increased under the more flexible pricing regimes. Notably, the increase in consumer surplus is more than double the increase in profits (equivalent to 80% of firm profits under two-part tariffs), implying that the efficiency gains are mostly reaped by consumers, which may speak to the high degree of competition between the eight firms in this market. This is consistent with results on third-degree price discrimination under competition (such as Holmes, 1989), showing price discrimination is consumer welfare enhancing only if quantities are increased.

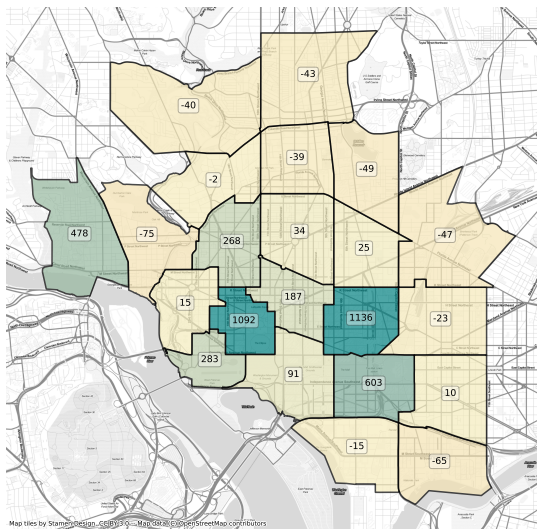
Notice, however, that the outcomes under the two flexible pricing regimes are nearly identical to each other. This is due to the relatively stable nature of vehicle stocks in this empirical setting, so that the incentive to respond to changes in the stocks occurs relatively rarely.



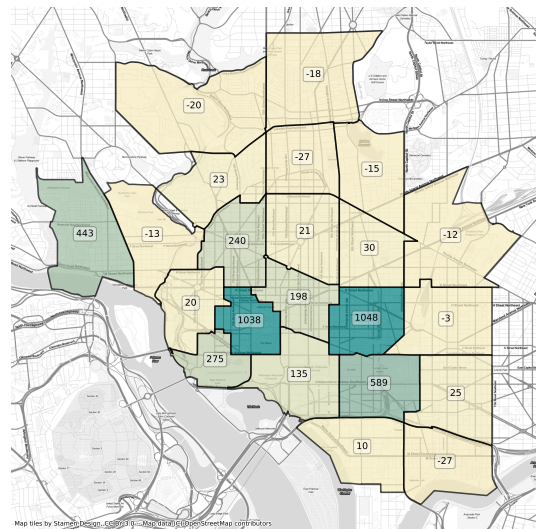
(a) Profits by destination



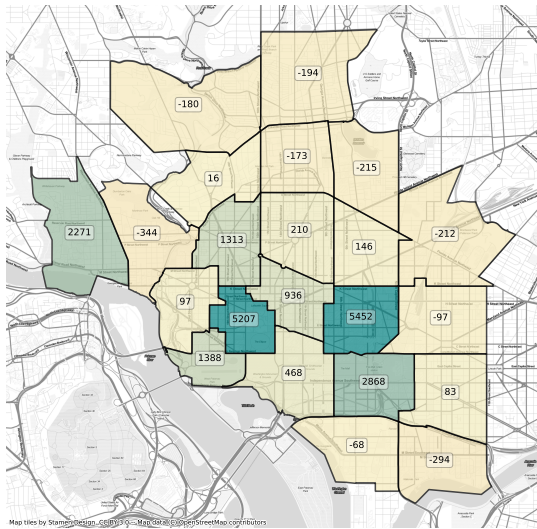
(b) Profits by origin



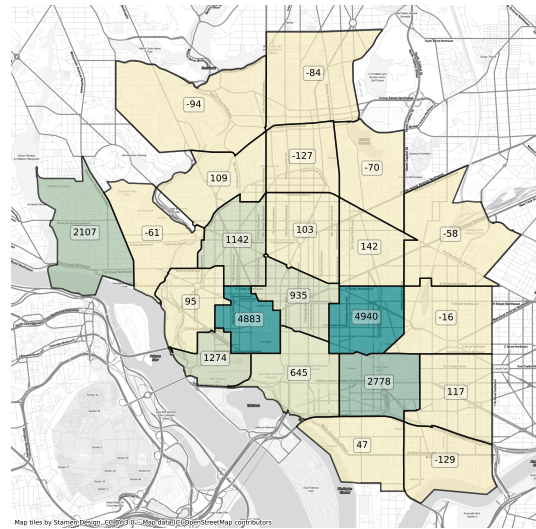
(c) Quantities by destination



(d) Quantities by origin



(e) Consumer surplus by destination



(f) Consumer surplus by origin

Figure 8: Daily outcomes: Flexible prices vs. two-part tariffs

Furthermore, the average stocks by location, time of day, and firm strongly coincide across the two pricing regimes, as shown in Figure A.5. In addition, conditional on the initial vehicle distribution, a firm’s prices charged under the two regimes for the same trip are also very closely correlated (see the example for Jump in Figure A.6). This suggests that this form of “surge pricing” has a limited role to play in this environment, possibly due to the lack of aggregate shocks in the model. The fact that the stock responsive prices actually perform worse despite being more flexible is attributable to some level of error inherent in the neural networks used for approximation. For the rest of this section, I thus focus on the flexible, but stock invariant prices.

Recall that the uniform price mandate effectively creates cross-subsidization across groups of consumers. In the context of transportation, this may result in certain areas facing higher prices or lower vehicle availability of vehicles when the pricing restrictions are lifted. To evaluate this, the comparison of outcomes under flexible prices and two-part tariffs are decomposed spatially by either trip origins or destinations in Figure 8.

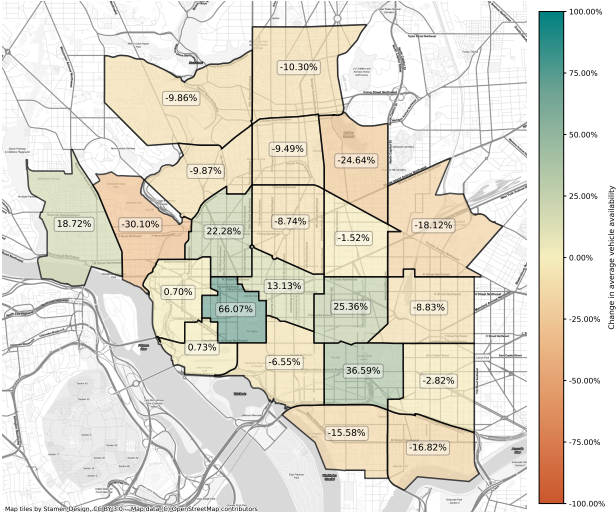


Figure 9: Flexible prices vs. two-part tariffs: vehicle availability

It is immediately clear that the changes in all three outcomes are highly concentrated in the downtown areas, particularly the central business district, as well as in the Georgetown University campus. The changes in vehicle availability shown in Figure 9 reflect the same pattern, suggesting that the flexible prices allow the stock of vehicles to be better managed, so that the increase in vehicle availability (which is analogous to an increase in product quality) explain the increases in consumer welfare, part of which is captured by the firms as increased profits.

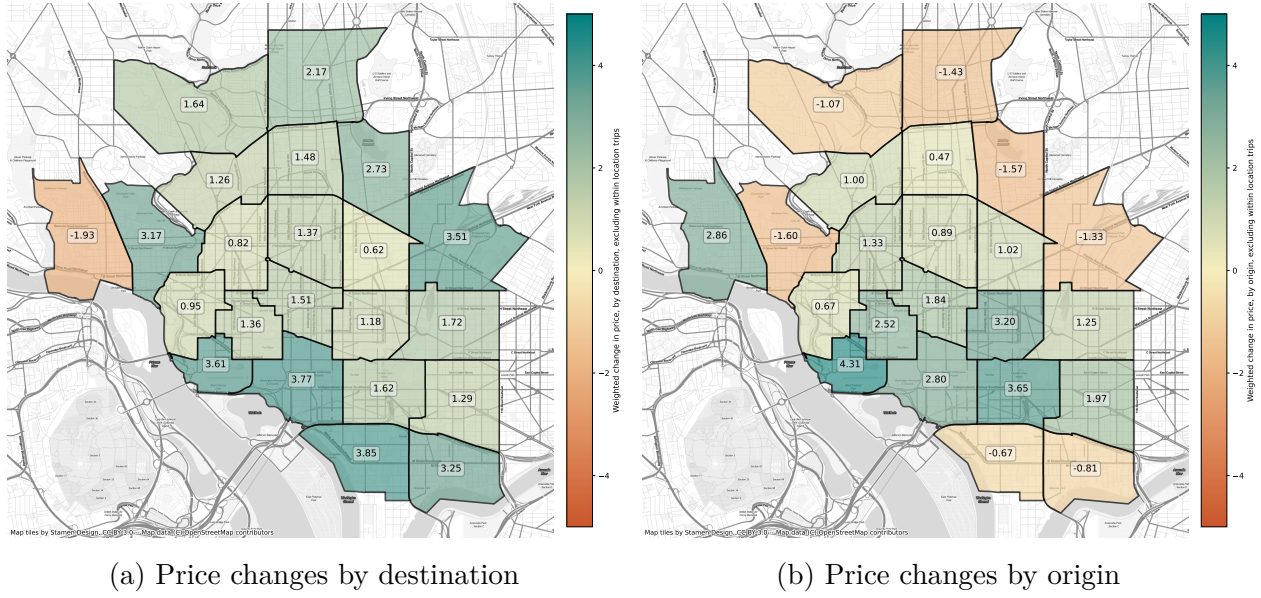


Figure 10: Quantity-weighted changes in cross-location prices

Note: Prices are weighted by the cross-location trip quantities (44% of all trips) under two-part tariffs.

Crucially, these increases in consumer welfare in downtown come at a fairly modest cost to the consumer welfare of the peripheral areas, despite significant decreases in the availability of vehicles. This can be explained by the change in cross-location prices depicted in Figure 10: prices for trips *originating* in the peripheral areas but heading elsewhere are *lowered* or only modestly increased. This acts as a countervailing effect to the reduction in supply, which addresses some of the policymakers' concern with equity of access to transportation.

To further analyze the distributional consequences of the policy, I regress the average welfare changes for each trip against the household income levels (in tens of thousands of dollars) for those same trips according to the Regional Travel Survey (see Section 3.2), as well as the trip's distance traveled *away* from the center of the city, taken to be the White House (the eleventh cell in the grid in Figure 4). The results, presented in Table 8, show a very clear relation between the welfare effect of the price control and both income and distances. However as shown in Figure 11a the relationship with income is U-shaped, so that the relaxation of the price control policy benefits riders at both the low- and high-end of the income range. The redistributive effects of the price controls are neither strictly progressive nor regressive. Instead, the price controls appear to mainly benefit riders going away from the city center (see Figure 11b), since they are reducing the vehicle supply in the most valuable areas. Thus, much like the uniform prices charged in New York City subways, the price controls mainly act as a subsidy to the peripheral areas, without achieving any clearly

progressive redistribution.

	Coef.	Std. Error	t	P > t	[0.025	0.975]
Intercept	1.7253	0.668	2.583	0.010	0.412	3.038
Income	-0.2194	0.080	-2.731	0.007	-0.377	-0.061
Income Squared	0.0067	0.002	2.914	0.004	0.002	0.011
Distance from center	-0.2653	0.065	-4.101	0.000	-0.392	-0.138

Table 8: Regression of average welfare change

One caveat of these results are the assumption which holds the initial distribution of vehicles fixed across the counterfactuals. While this is indeed a strong assumption, it can be justified under the current policies which regulate the initial distribution of vehicles (see Section 3). These regulations would appear to suffice in maintaining sufficient vehicle availability across locations, while the uniform pricing regulation may be dispensed with to allow the system to better maintain the optimal balance of vehicles. From this perspective, the flexible prices may actually complement regulations on the distribution of vehicles, as they can maintain the socially desirable distribution of vehicles (mandated by the regulator) at a lower cost by reducing the amount of manual relocation required.

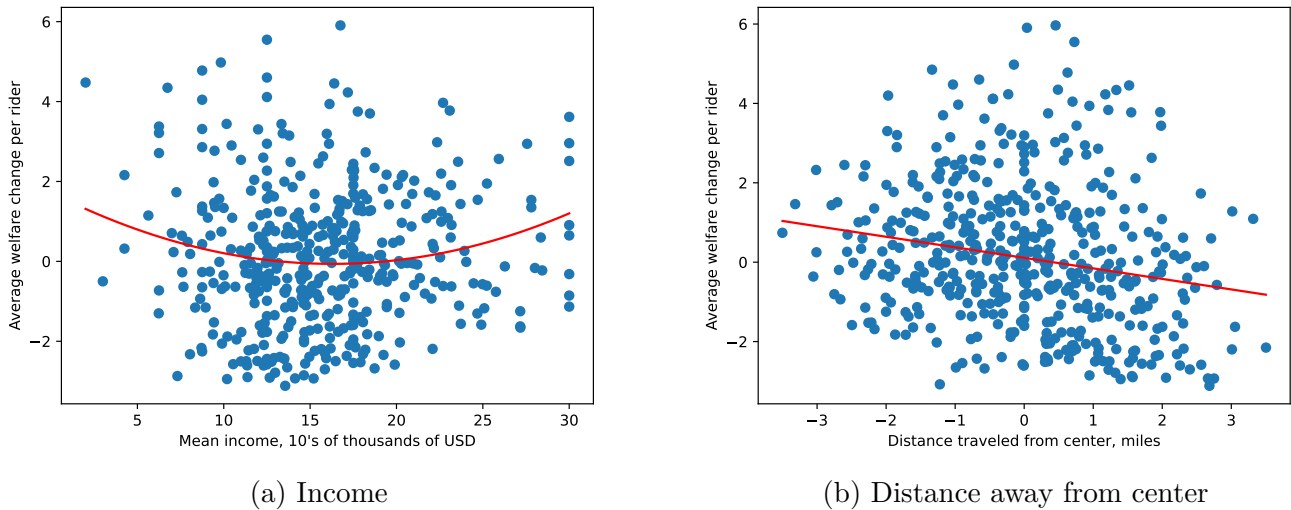


Figure 11: Decomposing welfare changes by income and distance

7 Conclusion

Price controls have a long history in transportation policy, ostensibly with the objective of effecting progressive income redistribution. Yet little is known about the kind of redistribution these policies actually cause, and how much economic efficiency must be given up by distorting prices away from their market determined levels. Such policies have seen renewed interest with the introduction of novel forms of transportation, most notably shared vehicle platforms. This paper sheds light on this policy question.

I find that the price controls impose a substantial harm on the economic efficiency of the market. Relaxing the price constraints would yield an increase in industry profits of 34%, or about \$6,000 per day and a consumer surplus increase of \$14,700 per day with an increase in trips taken of about 41%. The redistributive effects of the policy do not appear to be progressive, but rather act as a subsidy to trips towards the periphery of the city. The purpose of such a subsidy is unclear, though it may serve as an incentive to reduce urban density.

In conclusion, this paper underscores the importance of conducting thorough assessments of price controls within the realm of transportation, considering the significant externalities that can arise from individual riders' behaviors. Furthermore, the methodological advancements presented in this study enable the computation of counterfactual prices while accommodating various constraints on their flexibility. As a result, this novel methodology holds potential for evaluating the effectiveness of alternative price control measures not only within transportation but also in broader contexts. By shedding light on the complexities surrounding price controls, this research contributes to a more comprehensive understanding of their implications and opens avenues for further exploration and policy refinement.

References

- Adams, Brian and Kevin R. Williams (Feb. 2019). “Zone Pricing in Retail Oligopoly.” In: *American Economic Journal: Microeconomics* 11.1, pp. 124–56. DOI: 10.1257/mic.20170130. URL: <https://www.aeaweb.org/articles?id=10.1257/mic.20170130>.
- Anas, Alex, Richard Arnott, and Kenneth A Small (1998). “Urban spatial structure.” In: *Journal of economic literature* 36.3, pp. 1426–1464.
- Arrow, Kenneth J (1951). “An extension of the basic theorems of classical welfare economics.” In: *Proceedings of the second Berkeley symposium on mathematical statistics and probability*. Vol. 2. University of California Press, pp. 507–533.
- Asker, John, Chaim Fershtman, and Ariel Pakes (2023). “The impact of artificial intelligence design on pricing.” In: *Journal of Economics & Management Strategy* n/a.n/a, pp. 1–29. DOI: 10.1111/jems.12516. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jems.12516>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jems.12516>.
- Azinovic, Marlon, Luca Gaegauf, and Simon Scheidegger (2022). “Deep equilibrium nets.” In: *International Economic Review* 63.4, pp. 1471–1525.
- Berry, Steven, Amit Gandhi, and Philip Haile (2013). “Connected substitutes and invertibility of demand.” In: *Econometrica* 81.5, pp. 2087–2111.
- Boadway, Robin and Michael Keen (2000). “Redistribution.” In: *Handbook of income distribution* 1, pp. 677–789.
- Bodéré, Pierre (2023). *Dynamic spatial competition in early education: An equilibrium analysis of the preschool market in Pennsylvania*. Tech. rep. Working paper.
- Brancaccio, Giulia et al. (2020). *Search Frictions and Efficiency in Decentralized Transport Markets*. Tech. rep. mimeo, Harvard University.
- Buchholz, Nicholas (Sept. 2021). “Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry.” In: *The Review of Economic Studies* 89.2, pp. 556–591. ISSN: 0034-6527. DOI: 10.1093/restud/rdab050. eprint: https://academic.oup.com/restud/article-pdf/89/2/556/42748180/rdab050_supplementary_data.pdf. URL: <https://doi.org/10.1093/restud/rdab050>.
- Buchholz, Nicholas et al. (May 2020). *The Value of Time: Evidence From Auctioned Cab Rides*. Working Paper 27087. National Bureau of Economic Research. DOI: 10.3386/w27087. URL: <http://www.nber.org/papers/w27087>.
- Calvano, Emilio et al. (2020). “Artificial intelligence, algorithmic pricing, and collusion.” In: *American Economic Review* 110.10, pp. 3267–97.

- Castillo, Juan Camilo (2022). “Who benefits from surge pricing?” In: *Available at SSRN 3245533*.
- Cho, Sungjin and John Rust (Apr. 2010). “The Flat Rental Puzzle.” In: *The Review of Economic Studies* 77.2, pp. 560–594. ISSN: 0034-6527. DOI: 10.1111/j.1467-937X.2009.00556.x. eprint: <https://academic.oup.com/restud/article-pdf/77/2/560/18365587/77-2-560.pdf>. URL: <https://doi.org/10.1111/j.1467-937X.2009.00556.x>.
- Cohen, Peter et al. (2016). *Using big data to estimate consumer surplus: The case of uber*. Tech. rep. National Bureau of Economic Research.
- Conlon, Christopher T. and Julie Holland Mortimer (2013). “Demand estimation under incomplete product availability.” In: *American Economic Journal: Microeconomics* 5.4, pp. 1–30.
- Corts, Kenneth S. (1998). “Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment.” In: *The RAND Journal of Economics* 29.2, pp. 306–323. ISSN: 07416261. URL: <http://www.jstor.org/stable/2555890>.
- DDOT (2019). *TERMS AND CONDITIONS FOR THE PUBLIC RIGHT-OF-WAY OCCUPANCY PERMIT*. https://ddot.dc.gov/sites/default/files/dc/sites/ddot/page_content/attachments/Dockless%20Terms%20and%20Conditions%20-%20Phase%20II%20-%20Scooters%20-%2012.19.18.pdf. Accessed: 2019-11-18.
- (2020). *TERMS AND CONDITIONS FOR THE PUBLIC RIGHT-OF-WAY OCCUPANCY PERMIT*. https://ddot.dc.gov/sites/default/files/dc/sites/ddot/page_content/attachments/2019.11.6%20Shared%20dockless%202020%20Terms%20and%20Conditions%20scooter.pdf. Accessed: 2019-11-18.
- DellaVigna, Stefano and Matthew Gentzkow (June 2019). “Uniform Pricing in U.S. Retail Chains*.” In: *The Quarterly Journal of Economics* 134.4, pp. 2011–2084. ISSN: 0033-5533. DOI: 10.1093/qje/qjz019. eprint: <https://academic.oup.com/qje/article-pdf/134/4/2011/32666254/qjz019.pdf>. URL: <https://doi.org/10.1093/qje/qjz019>.
- Erkan, Hale et al. (2022). “Why so many scooters? A policy analysis.” In: *A policy analysis (May 5, 2022)*.
- Galbraith, John Kenneth (1980). *A theory of price control*. Harvard University Press.
- Gilligan, Thomas W, William J Marshall, and Barry R Weingast (1990). “The economic incidence of the Interstate Commerce Act of 1887: A theoretical and empirical analysis of the short-haul pricing constraint.” In: *The Rand Journal of Economics*, pp. 189–210.
- Hayek, Friedrich A. (1945). “The Use of Knowledge in Society.” In: *American Economic Review* 35.4, pp. 519–530.

- Holmes, Thomas J. (1989). “The Effects of Third-Degree Price Discrimination in Oligopoly.” In: *The American Economic Review* 79.1, pp. 244–250. ISSN: 00028282. URL: <http://www.jstor.org/stable/1804785>.
- Igami, Mitsuru (2020). “Artificial intelligence as structural estimation: Deep Blue, Bonanza, and AlphaGo.” In: *The Econometrics Journal* 23.3, S1–S24.
- Iskhakov, Fedor, John Rust, and Bertel Schjerning (2020). “Machine learning and structural econometrics: contrasts and synergies.” In: *The Econometrics Journal* 23.3, S81–S124.
- Kahou, Mahdi Ebrahimi et al. (2021). *Exploiting symmetry in high-dimensional dynamic programming*. Tech. rep. National Bureau of Economic Research.
- Lu, Tao, Chung-Yee Lee, and Loo-Hay Lee (2020). “Coordinating pricing and empty container repositioning in two-depot shipping systems.” In: *Transportation Science* 54.6, pp. 1697–1713.
- Miravete, Eugenio J., Katja Seim, and Jeff Thurk (Feb. 2020). “One Markup to Rule Them All: Taxation by Liquor Pricing Regulation.” In: *American Economic Journal: Microeconomics* 12.1, pp. 1–41. DOI: 10.1257/mic.20180155. URL: <https://www.aeaweb.org/articles?id=10.1257/mic.20180155>.
- Mirrlees, James A (1971). “An exploration in the theory of optimum income taxation.” In: *The review of economic studies* 38.2, pp. 175–208.
- Mnih, Volodymyr et al. (2015). “Human-level control through deep reinforcement learning.” In: *Nature* 518.7540, pp. 529–533.
- Newey, Whitney K and Daniel McFadden (1994). “Large sample estimation and hypothesis testing.” In: *Handbook of econometrics* 4, pp. 2111–2245.
- Orbach, Barak Y. and Liran Einav (2007). “Uniform prices for differentiated goods: The case of the movie-theater industry.” In: *International Review of Law and Economics* 27.2, pp. 129–153. ISSN: 0144-8188. DOI: <https://doi.org/10.1016/j.irl.2007.06.002>. URL: <http://www.sciencedirect.com/science/article/pii/S0144818807000488>.
- Pakes, Ariel and Paul McGuire (2001). “Stochastic algorithms, symmetric Markov perfect equilibrium, and the ‘curse’ of dimensionality.” In: *Econometrica* 69.5, pp. 1261–1281.
- Pandey, Akshat and Aylin Caliskan (2021). “Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy’s Price Discrimination Algorithms.” In: *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 822–833.
- Rosaia, Nicola (2020). *Competing platforms and transport equilibrium: Evidence from New York City*. Tech. rep. mimeo, Harvard University.
- Rust, John (2000). *Parametric Policy Iteration: An Efficient Algorithm for Solving Multidimensional DP Problems?* Tech. rep. Mimeo, Yale University.

- Shiller, Ben and Joel Waldfogel (2011). “Music for a Song: An Empirical Look at Uniform Pricing and Its Alternatives.” In: *The Journal of Industrial Economics* 59.4, pp. 630–660. DOI: 10.1111/j.1467-6451.2011.00470.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-6451.2011.00470.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6451.2011.00470.x>.
- Silver, David et al. (2016). “Mastering the game of Go with deep neural networks and tree search.” In: *Nature* 529.7587, pp. 484–489.
- Sutton, Richard S and Andrew G Barto (2018). *Reinforcement learning: An introduction*. MIT press.
- Sutton, Richard S et al. (1999). “Policy gradient methods for reinforcement learning with function approximation.” In: *Advances in neural information processing systems* 12.
- Sweeting, Andrew (2013). “Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry.” In: *Econometrica* 81.5, pp. 1763–1803.
- Taussig, Frank W (1919). “Price-Fixing as seen by a Price-Fixer.” In: *The Quarterly Journal of Economics* 33.2, pp. 205–241.

Online Appendix to Squeezing more Juice out of Lime: A Novel High-dimensional Pricing Algorithm

Richard Faltings*

March 17, 2024

A Additional Content

A.1 Companion Notebook

Section 2.1 is reproduced in a publicly accessible notebook on Google Colab at the following link: <https://colab.research.google.com/drive/1cMJtJbQ1jfibjtclYfh32V4R7poTqqzp?usp=sharing>

This notebook allows anyone to easily reproduce the results in that section of the paper, and to test the algorithm under alternative parameters.

A.2 Optimal Firm Relocation

The following is a formalization of the firm’s relocation problem, which is not solved for in the counterfactuals. As discussed in Section 4.2, I model relocation of vehicles (and charging) as taking place overnight. Here, I further assume that the firm’s relocation policy takes the form of a fixed initial vehicle distribution in the morning period, denoted \tilde{Q}_f^* . This is motivated by some of the firms’ business model, which adopts a “gig economy” approach to charging and relocation. This lets workers sign up through a mobile application and find

*The University of Texas at Austin: richard.faltings@utexas.edu

vehicles needing to be charged. These firms generally require vehicles that are taken for charging to be redeployed the following morning at specific locations.¹ Given this business model, a firm would need to commit to its deployment plan so gig workers can plan their routes more efficiently.

Intuitively, the firm must then choose its initial deployment such as to balance the value of a particular distribution of vehicles with its expected cost of relocating vehicles. Both the value and the cost of relocation will depend on its pricing policy, which affects how its vehicles will move over the course of the day.

Formally, I model the choice of initial deployment \check{Q}_f^* as:

$$\check{Q}_f^* = \operatorname{argmax}_{\check{Q}_f} \mathbb{E}_{\check{Q}_f | p_f^*, \check{Q}_f} [V_f(\check{Q}_f, h = 1 | p_f^*) - C^*(\check{Q}_f, \bar{Q}_f | C^r)] \quad (\text{A.1})$$

Where \bar{Q}_f is the firm’s vehicle distribution at the end of the day, which is a random variable whose probability distribution depends both on the pricing policy and the initial vehicle distribution. $V_f(\check{Q}_f, h = 1 | p_f^*)$ is the firm’s value of having the given vehicle distribution at the start of the day, defined according to the pricing problem in 7. I let $C^*(Q', Q | C^r)$ be the cost of optimally relocating vehicles from a distribution Q to a distribution Q' , obtained from a canonical optimal transport problem with movement cost matrix C^r (also known as the earth mover’s distance). I also use this to define $V_R(Q)$, the firm’s value of ending the day with a particular vehicle distribution, which will require relocation:

$$V_R(Q) = C^*(\check{Q}_f^*, Q | C^r) \quad (\text{A.2})$$

A.3 Counterfactual Algorithm Details

The algorithm described here solves for the parametric pricing equilibrium described in Section 4.3. For each firm $f \in \mathcal{F}$, the algorithm computes the optimal parameters θ_f^* for the firm’s parametric pricing function $p_{\theta_f, p}(Q_f, h)$.

In addition to the iteration of the pricing function described in Equation 4 in Section 6.1, the full algorithm also makes use of a value function to reduce the variance of the estimated gradients. The intuition is that part of the expected gradient involves the profit of the firm earned at a particular state (Q_f, h) , as seen in Equation 2. The profit multiplies the gradient

¹See: <https://help.bird.co/hc/en-us/articles/360031785072-24-7-Nest-Availability>

on the probability of the realized demand. Thus the gradient effectively puts a greater weight on realizations of q with higher profits, making these states more likely. However, since the profits are non-negative, no states are ever *penalized* with a negative weight. Instead, we can change the weighting to an *advantage* which subtracts the expected profit from the current state. This has the effect of centering the weights, leading to lower variance in the estimated gradient and better convergence of the stochastic gradient ascent but does not modify the optimum of the problem. This closely follows the concept of Generalized Advantage Estimation of Schulman et al. (2015). The algorithm thus uses a parametric approximation $V_{\phi_f}(Q_f, h)$ for each firm $f \in \mathcal{F}$.

Recall Equation 3 describing the stochastic gradient used to update the parametric policy.

$$\tilde{\nabla}_{\theta} J(\theta) = \frac{1}{S} \sum_{s=1}^S \left[\sum_{t=0}^{T_s} \left(\beta^t (\nabla_{\theta} p_{\theta}) q_{s,t} + (\nabla_{\theta} p_{\theta}) \underbrace{\nabla_p \log g(q_{s,t}|p, Q)}_{\text{log probability gradient}} \underbrace{\sum_{t=0}^{T_s} ((p-c)' q_{s,t})}_{\approx V(Q_t, h_t)} \right) \right] \quad (\text{A.3})$$

The last term to be summed over is essentially the profit obtained from the sequence of demand starting from time t , which has state Q_t and h_t . In expectation, this sequence will yield exactly $V(Q_t, h_t)$, by definition of the value function itself as the continuation value from a particular state. However, there is significant variance involved in drawing the sequences of demand. This is a source of variance in the stochastic gradient itself. As described in Schulman et al. (2015) we can reduce the simulation variance by subtracting the value function from the simulated profit sequence. This doesn't change the value of the gradient in expectation, but reduces its variance. Intuitively, we can think of the sequence of profits as a weight, which is multiplied with the log probability gradient. Both of these are random variables, so the variance of their product will depend on both their covariances and the product of their expectations. By subtracting the value function from the profit weights, we reduce the expectation of the weights to 0, which also has the effect of reducing the variance of their product with the log probability gradient. Given the approximation $V_{\phi}(Q, h)$, we can thus use a modified gradient as follows:

$$\hat{\nabla}_{\theta} J(\theta) = \frac{1}{S} \sum_{s=1}^S \left[\sum_{t=0}^{T_s} \left(\beta^t (\nabla_{\theta} p_{\theta}) q_{s,t} + (\nabla_{\theta} p_{\theta}) \nabla_p \log g(q_{s,t}|p, Q) \left(\sum_{t=0}^{T_s} ((p-c)' q_{s,t}) - V_{\phi}(Q_t, h_t) \right) \right) \right] \quad (\text{A.4})$$

where we subtract $V_{\phi}(Q_t, h_t)$ from the profit sequence $\sum_{t=0}^{T_s} ((p-c)' q_{s,t})$.

At the same time, given different sequences of profit, we can also update the value function

approximation itself:

$$\hat{\nabla}_\phi J(\phi) = \frac{1}{S} \sum_{s=1}^S \left[\nabla_\phi \left(\sum_{t=0}^{T_s} ((p-c)'q_{s,t}) - V_\phi(Q_t, h_t) \right) \right] \quad (\text{A.5})$$

We can now define the full price optimization algorithm used to compute the counterfactual scenarios in Algorithm 1.

Algorithm 1 Stochastic Gradient Price Optimization

```

for all  $f \in \mathcal{F}$  do
   $\theta_f^0 \leftarrow$  reasonable random initialization
   $\phi_f^0 \leftarrow$  reasonable random initialization
end for
for  $k$  from 1 to  $K$  do
  for  $s$  from 1 to  $S$  do
    Simulate a full day with prices  $p_{\theta_f^k} \forall f \in \mathcal{F}$  (until  $T_s$  s.t.  $h_{T_s} = \bar{h}$ ) to obtain:
     $q_{s,f,t}, \forall f \in \mathcal{F}, t \leq T_s$ 
  end for
  for all  $f \in \mathcal{F}$  do
    Stochastic policy update:  $\theta_f^{k+1} = \theta_f^k + \alpha_\theta^k \hat{\nabla}_{\theta_f^k} J(\theta_f^k)$ 
    Stochastic value update:  $\phi_f^{k+1} = \phi_f^k + \alpha_\phi^k \hat{\nabla}_{\phi_f^k} J(\phi_f^k)$ 
  end for
end for

```

The parameters used for the various counterfactuals are displayed in Table A.1. All the value functions used were approximated by a neural network (one for each firm). Additionally the parameter update steps are performed with the Adam algorithm from Kingma and Ba (2014), which further tempers the step size for better convergence and uses a form of momentum to overcome local optima.

Tuning Parameter	Two-Part Tariff	Stock Invariant	Stock Responsive
Value learning rate α_V	5e-4	5e-4	5e-4
Price learning rate α_p	1e-3	5e-3	5e-6
Advantage Estimation Parameter λ	0.5	0.5	0.5
Batch size	150	150	150
Parallel threads	30	30	30
Batches	12000	18000	18200
Value neural network structure	(22; 64; 64; 64; 1)	(22; 64; 64; 64; 1)	(22; 64; 64; 64; 1)
Price neural network structure	N/A	N/A	(22; 64; 64; 64; 1936)

Neural networks are all MLPs with structure denoted as inputs, neurons in hidden layers, and outputs
 See Schulman et al. (2015) for an explanation of the λ parameter

Table A.1: Training Parameters

Additionally, the expected walking distance, which enters into the utility (see Equation 5), is re-computed in the counterfactuals by modeling the expected walking distance as a function of the number of scooters, with the following functional form:

$$\mathbb{E}(w_{\ell,f}|Q_{\ell,f};\omega) = \frac{\omega_{\ell,f}^1}{\omega_{\ell,f}^2 Q_{\ell,f} + \omega_{\ell,f}^3} + \omega_{\ell,f}^4 \quad (\text{A.6})$$

When both vehicles and consumers are uniformly distributed, the above formula can be analytically derived with $\omega_{\ell,f}^3 = 0$ and $\omega_{\ell,f}^4 = 0$. These additional parameters are added to account for non-normality, and all four parameters are fitted on the original data. An example of the fit can be found in Figure A.4.

All computation was carried out on the Texas Advanced Computing Center’s Stampede2 cluster, with runtimes of up to 48 hours depending on the pricing function specification.

B Additional Tables

Firm	Unlock fee (SQ)	Minute fee (SQ)	Unlock fee (CF)	Minute fee (CF)
Bird	\$1.00	\$0.39	\$0.95	\$0.08
Jump (Bike)	\$0.00	\$0.25	\$1.12	\$0.07
Jump (Scooter)	\$0.00	\$0.25	\$1.11	\$0.06
Lime	\$1.00	\$0.24	\$1.29	\$0.09
Lyft	\$1.00	\$0.24	\$1.06	\$0.08
Razor	\$1.00	\$0.24	\$1.15	\$0.05
Skip	\$1.00	\$0.25	\$1.22	\$0.09
Spin	\$0.00	\$0.29	\$1.16	\$0.05

SQ: Status-quo. CF: Counterfactual

Table A.2: Two-part tariff comparison

C Additional Figures

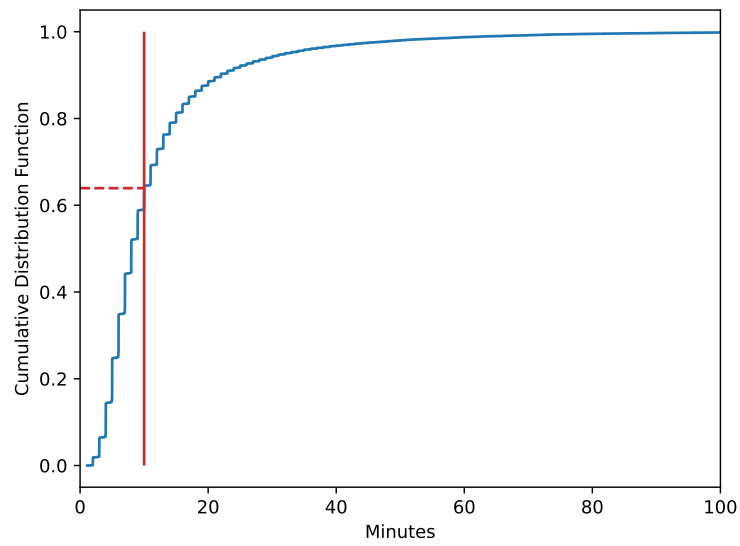


Figure A.1: CDF of trip durations for Jump scooters

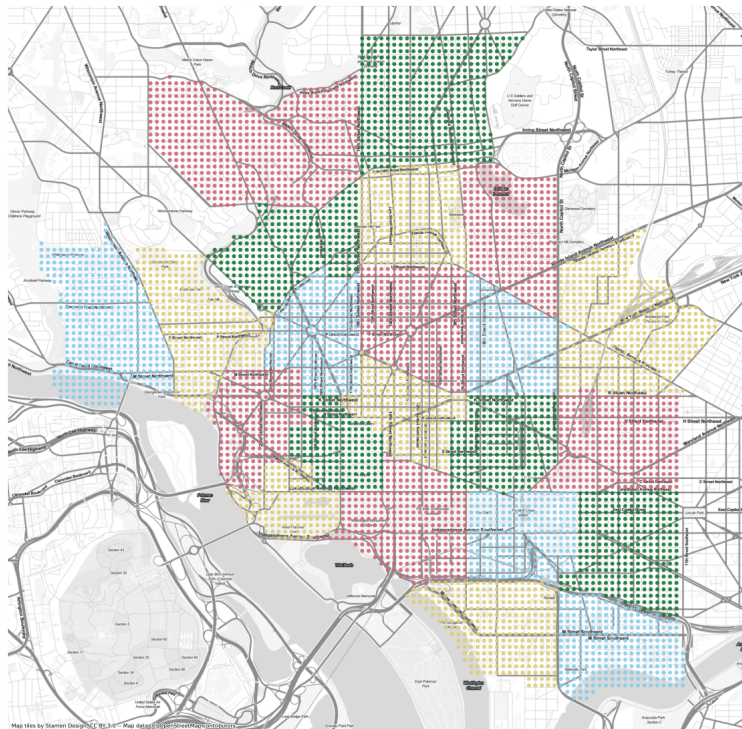


Figure A.2: Fine grid of points used to compute expected walking distances.

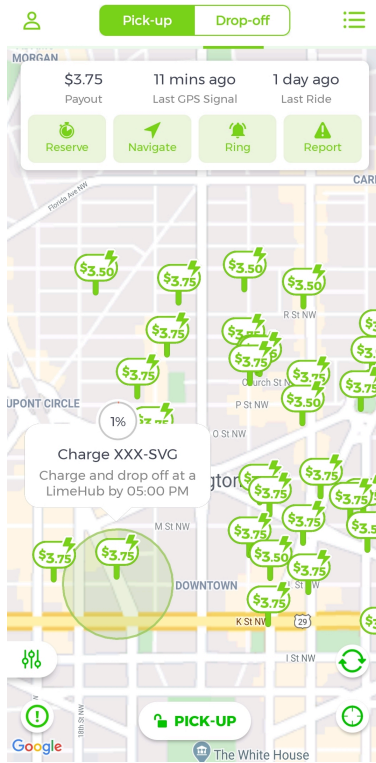


Figure A.3: Charging tasks and rewards for Lime scooters. Dated May 18th, 2020.

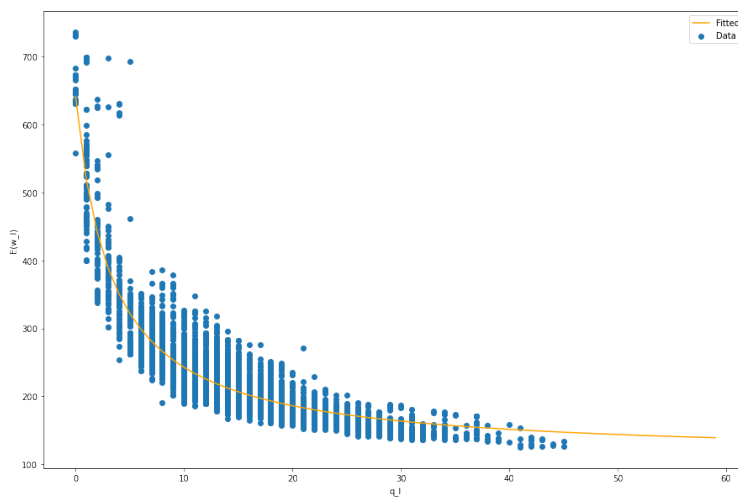


Figure A.4: Example of the fitted walking distance function

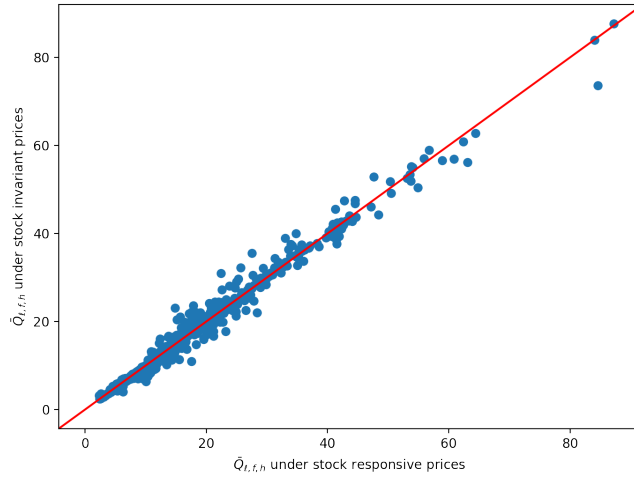


Figure A.5: Comparison of average vehicle stocks by firm, location, and hour of day

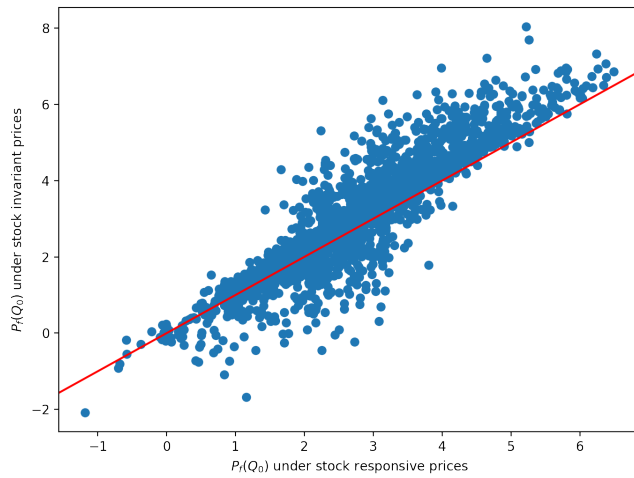
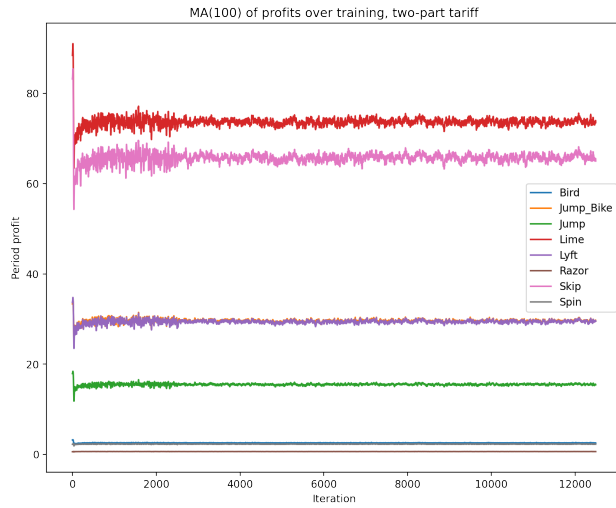
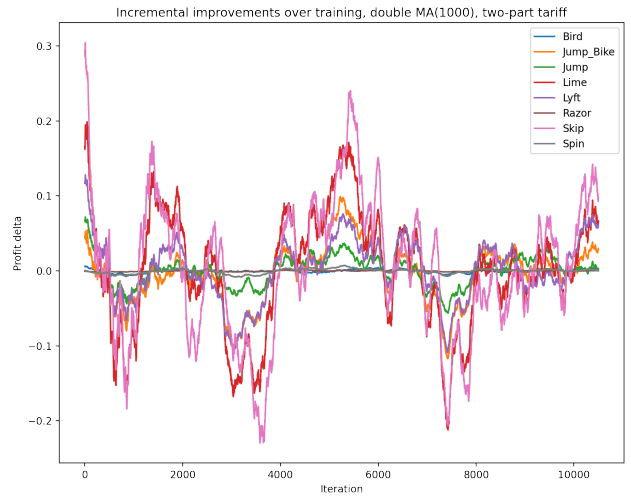


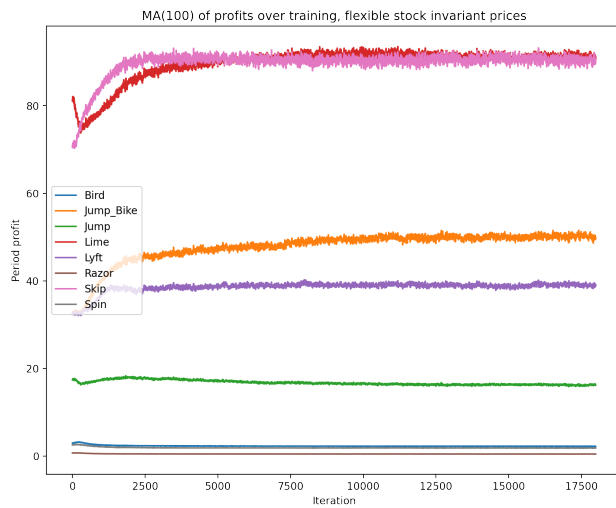
Figure A.6: Comparison of prices under initial vehicle distribution, Jump



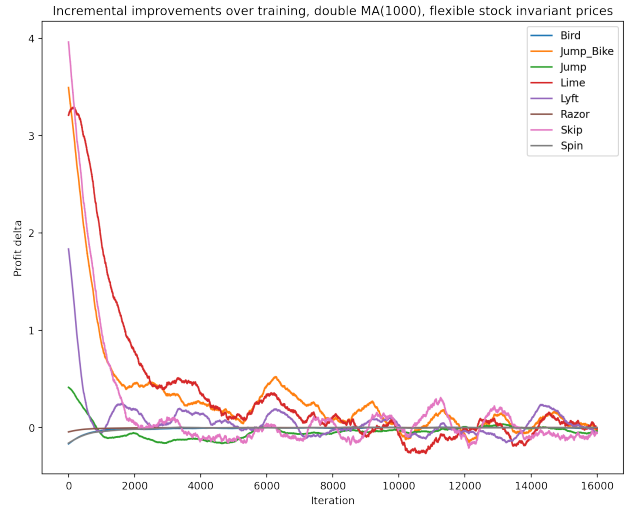
(a) Two-part tariff, profits over training



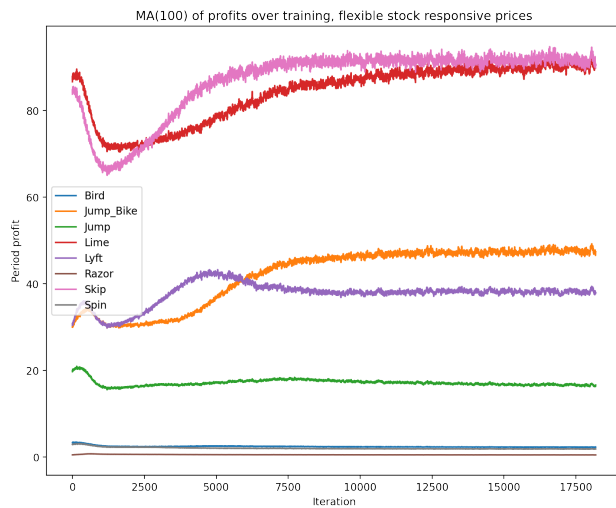
(b) Two-part tariff, training deltas



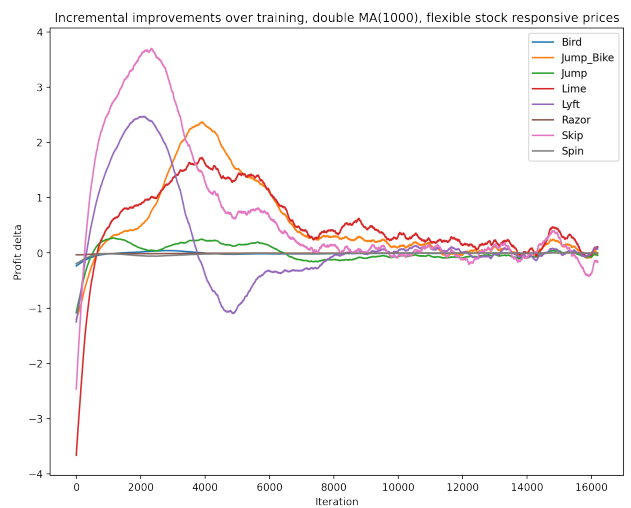
(c) Stock invariant prices, profits over training



(d) Stock invariant prices, training deltas



(e) Stock responsive prices, profits over training



(f) Stock responsive prices, training deltas

Figure A.7: Evolution of profits over training

References

- Kingma, Diederik P and Jimmy Ba (2014). “Adam: A method for stochastic optimization.” In: *arXiv preprint arXiv:1412.6980*.
- Schulman, John et al. (2015). “High-dimensional continuous control using generalized advantage estimation.” In: *arXiv preprint arXiv:1506.02438*.