

# Welfare Implications of Congestion Pricing: Evidence from *SFpark*

Pnina Feldman

Questrom School of Business, Boston University, pninaf@bu.edu

Jun Li

Stephen M. Ross School of Business, University of Michigan, junwli@umich.edu

Hsin-Tien Tsai

Department of Economics, National University of Singapore, ttsai@nus.edu.sg

**Problem definition:** Congestion pricing offers an appealing solution to urban parking problems—charging varying rates across time and space as a function of congestion may shift demand and improve allocation of limited resources. It aims to increase the accessibility of highly desired public goods and to reduce traffic caused by drivers who search for available parking spaces. At the same time, complex policies make it harder on consumers to make search-based decisions. We investigate the effect of congestion pricing on consumer and social welfare. **Academic/practical relevance:** This paper contributes to the theory and practice of the management of scarce resources in the public sector, where welfare is of particular interest. Methodologically, we contribute to the literature on structural estimation of dynamic search models. **Methodology:** Using data from the City of San Francisco, both before and after the implementation of a congestion pricing parking program, *SFpark*, we estimate the welfare implications of the policy. We use a two-stage dynamic search model to structurally estimate consumers’ search costs, distance disutilities, price sensitivities and trip valuations. **Results:** We find that congestion pricing increases consumer and social welfare in congested regions but may hurt welfare in uncongested regions, in which the focus should be on increasing utilization. Moreover, an unnecessarily complex congestion-pricing scheme makes it difficult for consumers to make search-based decisions. We find that a simpler pricing policy may yield higher welfare than a complex one. Lastly, compared to a policy that imposes limits on parking durations, congestion pricing increases social welfare by allocating the scarce resource to consumers who value it most. **Managerial implications:** The insights from *SFpark* offer important implications for local governments who consider alternatives for managing parking and congestion, and for public sector managers who evaluate the tradeoffs between approaches to manage public resources.

*Key words:* Congestion Pricing, Welfare, Dynamic Search Model, Public Sector, Traffic Management

---

## 1. Introduction

One of the challenges in managing public goods is to achieve an efficient allocation of resources while keeping utilization high. Without intervention by policy makers, individuals tend to overuse public goods and ignore the negative externalities they impose on others. This behavior leads

to congestion and other inefficient outcomes, a problem commonly referred to as the tragedy of commons. This problem is present in urban parking—affordable prices of public parking cause some users to overuse parking spaces without consideration of the negative impact to others. This behavior can induce urban transportation and other problems. As Shoup (2005) writes, “just as cattle compete in their search for scarce grass, drivers compete in their search for scarce curb parking spaces. Drivers waste time and fuel, congest traffic, and pollute the air while cruising for curb parking.” Summarizing sixteen studies, Shoup (2006) found that, on average, 30 percent of urban traffic was caused by drivers cruising to search for parking rather than driving to their desired destinations.

Congestion pricing is one solution to manage traffic congestion (Vickrey 1952). While Vickrey and others proposed the solution many decades ago, due to the technological challenges involved, it has only been put into practice recently. To implement congestion pricing, cities must install technologies such as cameras and sensors to track congestion levels frequently. In recent years, a few cities experimented with different variations of congestion pricing, including New York City’s PARK Smart (2008), San Francisco’s *SFpark* (2011), and Berkeley’s GoBerkeley (2012). With varying levels of pricing complexity, all programs reported increased accessibility and lower congestion (see reports for details of the various programs). While accessibility and decreased congestion are important, they are not all that matters. Social planners also care about maintaining high levels of utilization of public goods. Doing so is challenging because utilization induces congestion. Therefore, a good policy that strives to increase consumer and social welfare must strike a balance between utilization and congestion.

Using *SFpark*—the congestion pricing parking program implemented by the City of San Francisco—as a testbed, we wish to answer the following questions in this paper: (1) How does congestion pricing affect externalities caused by consumers, such as search? (2) Would congestion pricing lead to a more efficient allocation of public resources, and improve welfare via that allocation? and (3) What are the caveats of implementing congestion pricing?

To answer these questions, we model customers’ parking decisions using a two-stage dynamic structural model. In the first stage, a customer decides whether to drive and, if so, whether to park directly at a garage or to search for on-street parking. If a customer decides to search, in the second stage, she will make a dynamic decision of whether to park on street, continue to search or abandon searching. We estimate consumers’ search costs, distance disutilities and price sensitivities using availability and payment data from the *SFpark* program. We then use the estimates to quantify the effect of congestion pricing on consumer surplus, social welfare and search traffic.

We find several interesting results. First, our empirical analysis indicates that the effect of congestion pricing on consumer surplus depends on the level of congestion in a region—congestion

pricing may either increase or decrease welfare depending on the characteristics of the region we study. Congestion pricing increases welfare in popular regions with moderate to high congestion levels. However, it decreases welfare in less-congested areas. Second, even though it increases parking availability, a complex congestion pricing scheme makes it difficult for consumers to make search-based decisions (e.g., where to start, where to search, etc). and also induces search for lower prices. Interestingly, we find that a simpler three-tier pricing policy may increase welfare relative to a more complex policy, because consumers can be more informed and use this information to improve decisions. Finally, we compare the efficacy of congestion pricing to a policy that charges a fixed price, but sets time limits on parking (which was the policy in San Francisco prior to *SFpark*). We find that congestion pricing leads to higher social welfare but that the effect on consumer surplus is ambiguous.

What we learn from *SFpark* offers important lessons to local governments that consider alternative approaches to manage parking congestion. Congestion pricing is indeed an attractive approach to manage highly utilized public resources. It leads to higher welfare by allocating the resource to customers who value it the most. However, a good implementation of congestion pricing is nuanced. First, the level of congestion matters—congestion pricing may not work as well in uncongested areas. Second, the complexity of the pricing policy matters—if the pricing policy is very complex, congestion pricing may lead to inefficient search and decrease welfare. Therefore, a simpler policy is often more desirable than a complex one. Finally, our results highlight ways to manage scarce public resources better. Public sector managers often mitigate over-utilization by rationing capacity through usage limits or permits. We demonstrate that congestion pricing can be a more efficient approach. Congestion pricing accounts for heterogeneity in consumer demand through price discrimination, an aspect missing from capacity rationing levers. Of course, planners may have additional factors to consider, such as feasibility, cost, and equity concerns, when choosing a strategy. Nevertheless, our analysis offers quantifiable results and a generalizable methodology for public sector managers to better evaluate the tradeoffs involved in making such decisions.

## 2. Literature Review

This paper is related to three streams of literature: (1) dynamic pricing, price discrimination, and the effect of pricing on welfare; (2) public sector operations management; and (3) consumer demand modeling and structural estimation. We review each stream and discuss our contributions below.

**Dynamic Pricing, Price Discrimination and the effect of pricing on welfare.** Our paper contributes to the theory and practice of dynamic pricing. In the past several decades, dynamic pricing has been successfully applied in a number of industries, such as airlines, hotels and car rentals. More recently, additional industries, such as sports, concert planning and retail started

adopting these strategies (Shapiro and Drayer 2014, Xu et al. 2016, Tereyagoglu et al. 2016, Fisher et al. 2015, Moon et al. 2017). This line of research focuses primarily on profit/revenue maximization objectives, but a few empirical analyses examine the welfare effect of price discrimination. Leslie (2004) is one of the first studies to measure the welfare effect empirically. Unlike revenue or profit, welfare is not directly observable. Therefore, it is necessary to develop structural models to explicitly capture consumers' utility functions and decision processes. Using Broadway theater as an example, Leslie (2004) finds that while price discrimination leads to a 5% increase in firm profit, its impact on consumer surplus is negligible. Using airline data, Lazarev (2013) compares inter-temporal price discrimination to alternative pricing schemes (free resale, zero cancellation fees and third-degree price discrimination), and finds that the welfare effect is ambiguous and can be moderated by the mix of business vs. leisure travelers. We contribute to this line of research by empirically analyzing the welfare effect of dynamic pricing through structural models. We examine a dynamic pricing program implemented for city parking, where welfare is of particular interest. We offer insights on the conditions that affect the sign of welfare in the context of a public sector problem.

**Public Sector Operations Management.** There is a growing number of operations management papers studying the public sector. While a large body of this research focuses on healthcare, research is also burgeoning in education, public transportation, energy and utility, natural resource management and most recently the design of smart cities. Despite diverse contexts, a common theme that differentiates public versus private sector operations is the public sector's focus on societal outcomes rather than profitability. As a result, much emphasis has been placed on quality (e.g., Kc and Terwiesch 2009), congestion and utilization (e.g., Powell et al. 2012, Jaeker and Tucker 2017), accessibility (e.g., Kim et al. 2015, Gallien et al. 2016), and welfare and equity (e.g., Ashlagi and Shi 2016, Kok et al. 2016). The design of smart cities is receiving a lot of attention recently with advances in technology. Among the issues analyzed are sharing service integration (Qi et al. 2018), Electric vehicles and battery charging facilities (Mak et al. 2013 and Schneider et al. 2017) and bike sharing systems (Kabra et al. 2015). We contribute to this emerging topic by examining congestion pricing policies for city parking and their implications on welfare.

Stavins (2011) reviews and discusses two approaches to address the commons problem: the command and control approach (set usage limits) and the market-based approach (set prices to internalize the externalities). Despite the long-standing literature on optimal pricing (e.g., Vickrey 1952, Williamson 1966, Arnott and Inci 2006), empirical analyses of consumer reaction to dynamic pricing in public transportation are relatively scarce. Two recent studies (Pierce and Shoup 2013, Ottosson et al. 2013) estimate demand elasticity of changes of parking rates using regression approaches. Without explicit consumer decision models and structural estimation, however, they are unable to offer insights on the effects on welfare. We show that the market-based

approach leads to greater social welfare compared to a command and control policy, but that the effect on consumer surplus is ambiguous.

Finally, our work is related to studies of congestion in service operations, most of which analytically model the role of prices in regulating congestion in services, but recent papers also investigate the role of time limits (Tong and Rajagopalan 2014 and Feldman and Segev 2018). For extensive reviews see Hassin and Haviv (2003).

**Consumer Demand Modeling and Structural Estimation.** There is an increasing number of papers that use consumer choice models (Vulcano et al. 2010, Lederman et al. 2014, Kabra et al. 2015, Fisher et al. 2015), as well as models with dynamic decisions (Akşin et al. 2013, Li et al. 2014, Yu et al. 2016, Moon et al. 2017, Wang 2017, Emadi and Staats 2017). Our work is also closely related to literature on structural estimation of search models. This research estimates consumer search cost in different contexts, observing (De Los Santos et al. 2012, Honka 2014, Koulayev 2014, Chen and Yao 2016) or not observing (Hortacsu and Syverson 2004, Hong and Shum 2006, Kim et al. 2010) consumers' search paths. Similarly to the second set of papers, we also estimate search costs and other parameters without observing search paths. However, the data we use is more fine-grained and the search is multi-dimensional. Specifically, because the search is conducted on a two-dimensional map, it restricts the set of available options at every step. We embed a random walk with no immediate return to the dynamic search process. This enables us to address the challenge of dimensionality in estimation, while at the same time introduces randomness in the consumers' search process.

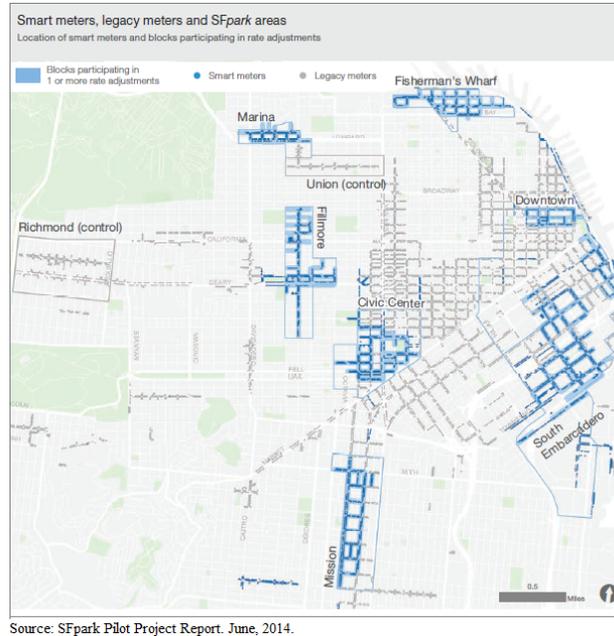
### 3. Background on the SFpark Program and Data Description

In this section, we introduce the SFpark program. We then describe the data used for this study and provide summary statistics for the periods before and after the implementation of the program.

#### 3.1. The SFpark program

The City of San Francisco implemented SFpark in 2011 to address urban parking problems via congestion pricing. Rather than charging a constant rate at all locations and at all times, the program adjusts parking rates according to demand. One of the challenges in implementing congestion pricing is that it requires constant monitoring of parking space utilization to adequately adjust prices. SFpark adopted several technologies, including parking sensors and smart meters, to track availability and measure utilization. The adoption of these technologies enabled SFpark to implement a data-driven parking pricing strategy. It also enabled researchers to conduct detailed analysis of consumer response to congestion pricing and its implications on welfare by using fine-grained data that were not previously available.

The San Francisco Municipal Transportation Agency (SFMTA) piloted the program in seven parking management regions (see Figure 1), which included 6,000 metered spaces amounting to roughly a quarter of the total metered parking spaces in San Francisco. The pilot started in August 2011 and ended in June 2013. The pilot was deemed successful: it illustrated the ability to reallocate demand, reduce congestion, and generate additional revenues. As a result, the program was rolled out to the entire city in late 2013.



**Figure 1** Smart Meters, Legacy Meters and SFpark Areas

With congestion pricing, *SFpark* adjusts hourly parking rates dynamically based on observed occupancy rates. The program divides each paid-parking day (Monday to Saturday) into three time windows: morning (9am-12pm), noon (12pm-3pm), and afternoon (3pm-6pm). Parking is free at other times and on Sundays. For each time window, *SFpark* uses the block-level average occupancy rate to determine the hourly rate for parking, where the occupancy rate is defined as the fraction of time that a block is occupied. *SFpark* started tracking occupancy rates of the piloted areas in April 2011, four months before the official start of the program. They used occupancy data during that period to determine price adjustment rules for the pilot, which started in August 2011. Before the implementation, parking was fixed at \$2 per hour for all blocks. After the implementation, *SFpark* raised a block's price by \$0.25/hour if the occupancy rate was above 80%, lowered it by \$0.25/hour if the occupancy rate was between 60% to 80%, and lowered it by \$0.50/hour if the occupancy rate was below 60%. *SFpark* also adjusted off-street parking prices (city-managed parking garages)

using similar rules.<sup>1</sup> Finally, *SFpark* set an upper- and a lower-bound for the hourly rate—the rate could not exceed \$6.00/hour or go below \$0.25/hour. As a result of these changes, parking rates varied by block, time of day, day of week, and month. Over the two-year pilot period, *SFpark* made ten on-street rate adjustments and eight off-street rate adjustments (i.e., every 8 to 12 weeks). All adjustments were announced on the program’s website at least seven days before the changes went into effect.

### 3.2. Data

We use three datasets provided by *SFpark*. Parking sensor data consist of hourly block-level occupancy rates from April 2011 until June 2013. After late 2012, however, the sensor data became incomplete due to battery failures and sensor outages. On-street meter payment data contain all parking transactions starting from the first quarter of 2011, and include start and end times, payment types and payment amounts. The meter payment data are more reliable than the parking sensor data because they are not subject to battery failures. However, meter payment data are not an accurate proxy for availability, because drivers may park for longer or shorter periods than paid for or may park illegally without paying. Hence, as long as the sensors were operating, the sensor dataset is a more accurate source for calculating occupancy rates. We therefore used the meter payment data to determine parking locations and durations, but did not to use it to infer occupancy rates. Off-street garage data contain usage data for publicly-owned parking garages. We observe transaction-level payment data at the same level of detail as meter payment data. The garage transaction data are not subject to illegal or under/overtime parking, because payment is determined based on actual parking time.

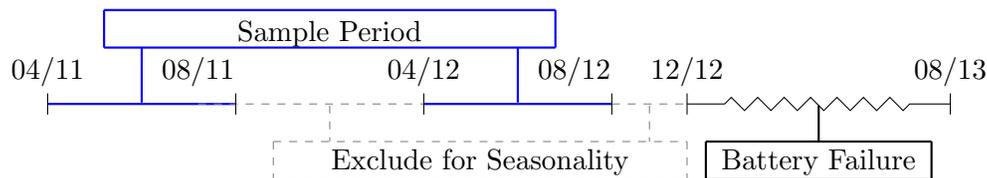


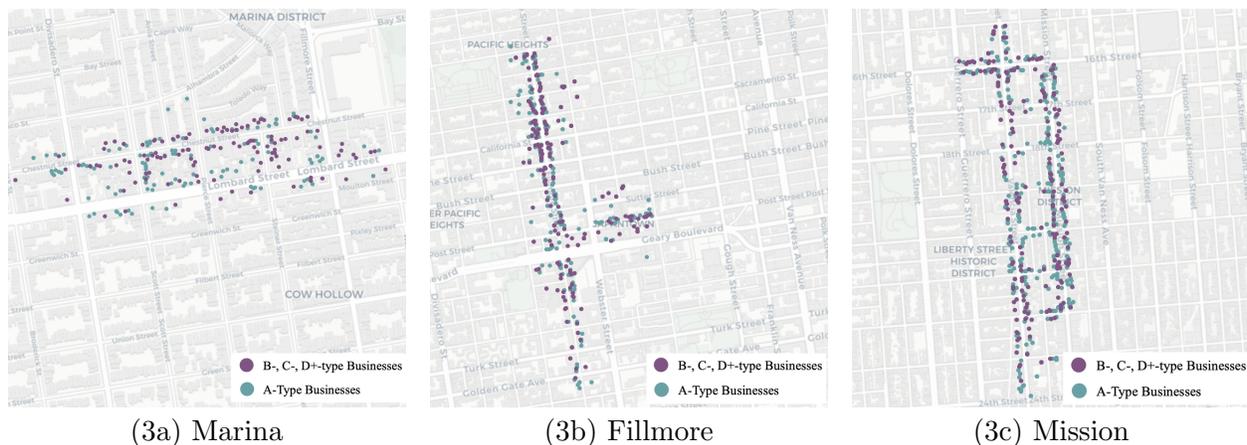
Figure 2 Timeline

Due to the increase in sensor failures starting in late 2012, we only use data from April 2011 to July 2012. To control for seasonality and to make fair comparisons between the before and after periods, we use data from the same months in both years: April to July 2011 (the before period),

<sup>1</sup> Before the implementation, the garage hourly parking rate ranged from \$2 to \$2.5. After the implementation, the hourly rate was raised by \$0.50 for blocks with occupancy rates above 80%, and lowered by \$0.50 for blocks with occupancy rates below 40%.

and April to July in 2012 (the after period) as shown in Figure 2.<sup>2</sup> The SFMTA extended the parking time limit in the pilot areas from 2 to 4 hours in late April, 2011. To make fair comparisons, we exclude the days in April 2011 in which the parking time limit was only 2 hours. We also exclude consumers who parked in a garage for more than 4 hours from the main analysis, but we do account for them for garage occupancy rate calculations. Among the seven piloted regions, we focus on the regions that are relatively more isolated from others: Fillmore, Marina, and Mission.<sup>3</sup>

In addition to the *SFpark* data, we also use the Infogroup US Historical Business Data in 2011 and 2012. The Infogroup Data provide the name, street address, and employee size of each registered business. We calculate the total numbers of businesses of different sizes in each parking block and merge them with the *SFpark* data. For parking blocks analyzed in our study, we identify a total of 2,051 businesses in 2011 and 2,216 businesses in 2012. Figure 3 shows the distribution of business establishments in Fillmore, Marina, and Mission in 2011. We use the number of businesses as an input in the destination model to generate parking demand. We allow businesses of different sizes to have heterogeneous impacts on the ideal location demand distribution. Specifically, the Infogroup data code the number of employees in a business as: A (1 to 4 employees), B (5 to 9 employees), C (10 to 19 employees), D (20 to 49 employees), etc. In the districts under study, the distribution of A-, B-, C-, and D-type businesses are 68.55%, 15.73%, 10.34%, 4.09% in 2011, and 69.17%, 14.87%, 10.78%, 3.83% in 2012. Larger businesses (i.e., with more than 50 employees) account for less than 1.5% of our data for both 2011 and 2012. We thus group them with type D and denote the combined group as D+.



**Figure 3 Business Density in 2011**

<sup>2</sup> Even during these periods, there were some occasional meter failures. We exclude these from the calculation of occupancy rates.

<sup>3</sup> Fisherman's Wharf is also relatively isolated. However, it is primarily a tourist destination. Since tourists might not have much knowledge of the *SFpark* program, they may make decisions differently and we exclude it from our study.

Table 1 presents the before and after summary statistics of the hourly parking rates, occupancy rates, and the number of businesses for the three regions. Consistent with SFpark guidelines, we use average occupancy rates to divide the parking blocks to high (average occupancy rates above 80%), medium (60% and 80% occupancy), and low-utilization (below 60% occupancy). Table 1 shows that after the implementation of congestion pricing, the mean parking rate increased by around 150% in high-utilization blocks, and decreased by between 40% - 70% in low-utilization blocks in Marina and Fillmore. In Mission, there were no high utilization blocks in the before period, and the parking rates in the low utilization blocks decreased slightly.<sup>4</sup> As expected, the average occupancy rate in low-utilization blocks increased while the average occupancy rate in high- and medium-utilization blocks decreased. This provides evidence of shifts in demand as a response to congestion pricing. We also find that high utilization blocks have a higher average number of businesses.

Figure 4 shows variations of parking rates and changes in occupancy rates between blocks in the after period via street map snapshots in Fillmore. The figure illustrates that prices vary substantially among blocks—a driver could save \$2 to \$3 per hour by driving one or two additional blocks. The figure also shows that congestion levels are more evenly distributed during the after period than the before period.

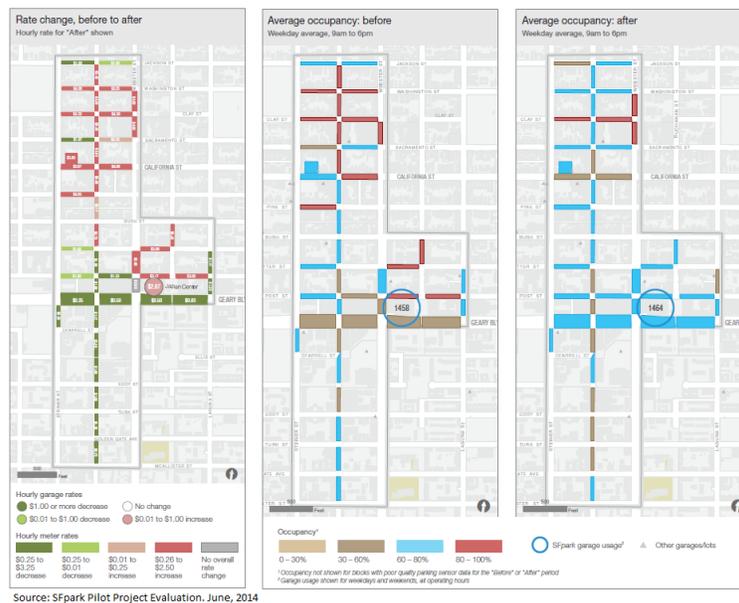


Figure 4 Snapshots of Fillmore

<sup>4</sup>Some parking spaces in Mission were blocked due to construction in March 2012. This induced higher occupancy rates in this area. For fair comparisons, we treat these blocks as available in welfare and counter-factual analyses.

**Table 1 Summary Statistics**

	Marina					Fillmore					Mission				
	Total	High	Mid	Low	Garage	Total	High	Mid	Low	Garage	Total	High	Mid	Low	Garage
Rate—Before	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	2.50 (0.00)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	(-)	2.00 (0.00)	2.00 (0.00)	2.25 (0.25)
Rate—After	2.50 (0.97)	3.24 (0.45)	2.61 (0.68)	1.17 (0.73)	2.33 (0.47)	2.27 (0.93)	3.29 (0.27)	2.24 (0.75)	1.31 (0.68)	2.33 (0.47)	2.32 (0.78)	(-)	2.53 (0.65)	1.92 (0.85)	2.00 (0.41)
Occupancy—Before	0.74 (0.16)	0.83 (0.11)	0.74 (0.13)	0.58 (0.16)	0.13 (0.06)	0.70 (0.19)	0.84 (0.15)	0.69 (0.16)	0.57 (0.17)	0.26 (0.13)	0.62 (0.18)	(-)	0.63 (0.17)	0.58 (0.18)	0.12 (0.06)
Occupancy—After	0.73 (0.16)	0.81 (0.12)	0.75 (0.13)	0.56 (0.15)	0.19 (0.08)	0.69 (0.20)	0.82 (0.16)	0.70 (0.16)	0.56 (0.21)	0.23 (0.14)	0.66 (0.18)	(-)	0.68 (0.18)	0.63 (0.18)	0.12 (0.05)
No. of A-Type Businesses—Before	8.68 (7.39)	9.17 (11.77)	7.22 (5.26)	11.25 (2.63)	(-)	16.07 (36.76)	39.90 (73.86)	9.80 (10.26)	7.90 (7.65)	(-)	20.72 (19.86)	(-)	25.89 (21.07)	7.43 (5.71)	(-)
No. of A-Type Businesses—After	9.84 (9.70)	12.17 (16.40)	7.67 (5.36)	11.25 (3.40)	(-)	16.93 (39.51)	42.60 (79.68)	10.32 (10.07)	7.80 (8.12)	(-)	22.08 (21.52)	(-)	27.22 (23.23)	8.86 (6.36)	(-)
No. of B-Type Businesses—Before	2.58 (3.32)	4.67 (5.43)	1.67 (1.22)	1.50 (0.58)	(-)	2.82 (3.21)	4.70 (4.47)	2.56 (2.96)	1.60 (1.17)	(-)	4.36 (3.41)	(-)	5.33 (3.38)	1.86 (2.04)	(-)
No. of B-Type Businesses—After	2.58 (3.13)	4.33 (5.16)	2.11 (1.27)	1.00 (0.00)	(-)	2.87 (3.20)	4.40 (4.20)	2.76 (3.17)	1.60 (1.07)	(-)	4.48 (3.66)	(-)	5.44 (3.75)	2.00 (2.00)	(-)
No. of C-Type Businesses—Before	2.32 (2.67)	3.17 (3.92)	2.11 (2.20)	1.50 (1.29)	(-)	2.04 (2.41)	2.20 (2.20)	2.00 (2.53)	2.00 (2.54)	(-)	3.36 (2.61)	(-)	4.17 (2.60)	1.29 (1.11)	(-)
No. of C-Type Businesses—After	2.74 (2.84)	3.83 (4.31)	2.44 (2.19)	1.75 (0.96)	(-)	2.24 (2.40)	2.30 (2.16)	2.24 (2.55)	2.20 (2.49)	(-)	3.84 (2.67)	(-)	4.78 (2.51)	1.43 (1.13)	(-)
No. of D-Type Businesses—Before	1.68 (1.73)	1.67 (2.07)	1.67 (1.87)	1.75 (1.26)	(-)	2.00 (4.37)	3.30 (6.43)	1.20 (1.53)	2.70 (6.46)	(-)	1.84 (1.70)	(-)	2.06 (1.70)	1.29 (1.70)	(-)
No. of D-Type Businesses—After	2.21 (2.20)	2.50 (3.21)	1.89 (1.90)	2.50 (1.29)	(-)	2.42 (5.02)	3.80 (7.36)	1.56 (1.94)	3.20 (7.36)	(-)	2.28 (1.97)	(-)	2.83 (1.89)	0.86 (1.46)	(-)
No. of Blocks	19	6	9	4	1.00	45	10	25	10	1	25	0	18	7	2
No. of Space	329	100	153	76	205	739	123	453	163	920	861	0	616	245	448

\* Standard deviations are in parentheses. When calculating the occupancy rate, we excluded non-operational hours for parking spaces when applicable, for example, peak-time tow away zones. “Before” refers to our sample period before congestion pricing: April to July in year 2011, “after” refers to our sample period after congestion pricing: April to July in year 2012. High, medium and low utilization blocks are defined using average occupancy rate in year 2011 (i.e., before SFpark) greater than 80%, between 60% and 80%, and below 60%, respectively. A-type business has 1–4 employees, B-type business has 5–9 employees, C-type business has 10–19 employees, and D+-type business has more than 20 employees.

## 4. Model

There are  $M_r$  consumers who are interested in visiting region  $r$  (i.e., Fillmore, Marina, Mission). In the baseline model, we set the market size of each region,  $M_r$ , to be twice the average number of drivers in the before period, which yields a market size of 700 in Marina, 1,135 in Fillmore and 1,420 in Mission. We later perform robustness checks to make sure that our results are insensitive to these market sizes.

We specify the decision process of a consumer  $i$ , whose trip destination is at block  $b_i^*$ , and is interested in parking for a duration of  $h_i$  hours. Block  $b_i^*$  is the ideal parking location for consumer  $i$ , if parking is free and available. Although we do not observe consumers’ ideal locations, we estimate the distribution of ideal locations of consumers as a function of business densities over the set of blocks,  $B_r$ , in region  $r$ : The fraction of consumers whose ideal location is block  $b$  is  $\omega_r(b)$ , and  $\sum_{b \in B_r} \omega_r(b) = 1$ . Details of this specification are in Section 5.1.

In our model, ideal locations and trip durations are determined endogenously and do not change once a customer has parked. While there may be situations in which a customer is willing to change her destination or duration based on congestion levels and parking rates, these two variables are largely determined by the purpose of the trip. For example, the destination of a consumer who plans to buy an iPhone is the Apple store, and the duration of the trip is determined by the expected time it takes to shop and purchase an iPhone.

We model consumers' driving and parking behavior as a series of decisions. We assume that each consumer chooses among three options: (1) drive to the region and search for on-street parking, (2) drive to the region but park directly in the public garage without searching,<sup>5</sup> or (3) choose an outside option, which includes staying at home, using other modes of transportation, or parking elsewhere.<sup>6</sup> A consumer chooses the option that gives her the highest expected utility. Without loss of generality, we normalize the mean utility of the outside option to zero. Customer  $i$ 's utility of the outside option is  $u_{i0} = \epsilon_{i0} \equiv V_i^o$ , where  $\epsilon_{i0}$  is an idiosyncratic shock to the outside utility of customer  $i$ , which follows a normal distribution with mean 0 and standard deviation  $\sigma$ .

Consumer  $i$  obtains a mean trip value,  $v_{irt}$ , from driving to the region relative to the outside option, irrespective of whether she parks at the garage directly or searches for on-street parking. We let the mean utility of driving be a linear function of the duration of parking. Specifically,  $v_{irt} = \alpha + \beta X_{rt} h_i$ , where  $X_{rt}$  contains the intercept and dummy variables indicating the time of day, day of week, and month. Even though all parameters are region specific, we omit the subscript  $r$  for parameters for brevity. We follow the literature on discretionary services and allow the mean utility of driving to increase with trip duration (e.g., Anand et al. 2011, Feldman and Segev 2018) and assume that this function is linear.<sup>7</sup> Keeping everything else equal, longer trips provide greater utility to the customer and this increase in utility offsets the parking costs, which also increase with trip duration. Whether a consumer chooses to drive depends on her mean utility but also on her costs and the utility shock, which taken in sum yields a non-trivial travel decision.

Customer  $i$  who decides to drive and park directly at the garage obtains value  $v_{irt} + \epsilon_{ig}$  from the trip, where  $\epsilon_{ig}$  is an idiosyncratic shock to her utility from parking at the garage, and follows the same distribution as  $\epsilon_{i0}$ . The customer also incurs costs for walking from the garage to her ideal location,  $b_i^*$ , and the garage parking fee, which is based on the hourly parking rate at the garage,  $P_{bg}$ , and the parking duration,  $h_i$ . Specifically, customer  $i$ 's utility for parking at the garage is:  $u_{ig} = \alpha + \beta X_{rt} h_i + \epsilon_{ig} - \eta_i d(b_i^*, b_g) - \theta_i P_{bg} h_i \equiv V_i^{\text{garage}}$ , where  $\eta_i$  is customer  $i$ 's cost of walking one

<sup>5</sup> In most regions we study, there is one public garage operated by the city. If there are multiple parking garages, a customer chooses the one that provides the highest utility based on her destination. We do not have data from private garages, and therefore include parking at a private garage as part of the outside option.

<sup>6</sup> We do not explicitly model the choice of a departure time (e.g., morning or afternoon, Monday or Tuesday). However, such inter-temporal shifts in demand are incorporated implicitly, to some extent, through the outside option. For instance, high prices on Tuesday afternoons result in fewer consumers driving to their destination and more consumers choosing the outside option.

<sup>7</sup> We have also analyzed alternative model specifications, in which  $v_{irt}$  is a function of  $X_{rt}$  only, a function of  $h_i$  only, or a linear additive function of  $X_{rt}$  and  $h_i$ . None of these specifications fit our empirical observations as well. In addition, to verify that a linear function of durations is reasonable, we estimate a model in which the trip valuation is a quadratic function of the trip duration. We find that in all regions, the estimated coefficients of the quadratic terms are all very close to and not statistically different from zero, suggesting that the linear model is a reasonable proxy for customer utility.

block,  $d(b_i^*, b_g)$  is the distance from the garage to her destination in blocks,  $b_i^*$ ,<sup>8</sup> and  $\theta_i$  is customer  $i$ 's price sensitivity. We assume that there is always an available parking space at the garage.<sup>9</sup>

Finally, a consumer who chooses to drive to the region and search for on-street parking would either end up parking at a block that she finds available and affordable, or could eventually decide to abandon searching and either park at the garage or choose the outside option (e.g., forgo the trip or park elsewhere). If she parks on-street, she will obtain value  $v_{irt} + \epsilon_{is}$ , where  $\epsilon_{is}$  is an idiosyncratic shock to consumer  $i$ 's utility from on-street parking and follows the same distribution as  $\epsilon_{i0}$  and  $\epsilon_{ig}$ . Note that all utility shocks,  $\epsilon_{i0}$ ,  $\epsilon_{ig}$  and  $\epsilon_{is}$ , are observable to the consumer, but not to the researchers. As with garage parking, a consumer who parks on-street pays for parking, and incurs search costs and the cost of walking to her destination if she parks in another block.

#### 4.1. Dynamic Search Model

##### *States, Actions and Utilities*

We derive the model of search with congestion pricing, in which prices may vary across blocks and by time of day, day of week, and month. Fixed pricing is a special case of this model. On the  $k$ th search, consumer  $i$  arrives at block  $b_k$ ,  $k = 1, 2, 3, \dots$ . There are three actions,  $a$ , that she can choose from: continue to search ( $a = 0$ ), park at the current block if there is a spot available ( $a = 1$ ), or stop searching for on-street parking ( $a = 2$ ). She chooses the option that gives her the highest expected utility. The utility of each option depends on the following state variables, which are realized after consumer  $i$  arrives to the block:

- $b_k$ : the block that the consumer arrives at on the  $k$ th search;
- $P_{rt}b_k$ : the hourly parking price at block  $b_k$  in region  $r$  at time  $t$  on day  $d$ ;
- $A_{rt}b_k$ : the availability of block  $b_k$  in region  $r$  at time  $t$  on day  $d$ .  $A_{rt}b_k$  equals one if there is at least one parking spot available and zero otherwise;
- $\epsilon_{irt}b_k$ : the shock to the cost to search observed by consumer  $i$  at block  $b_k$  in region  $r$  at time  $t$  on day  $d$ .

Consumer  $i$ 's utility from choosing action  $a$  at block  $b_k$  is  $u_i(b_k, P_{rt}b_k, A_{rt}b_k, \epsilon_{irt}b_k; a)$ . We specify the utility from each action below:

<sup>8</sup> The distance between two blocks is calculated as the minimum number of blocks that one has to walk from one block to the other. We assume all blocks have the same length.

<sup>9</sup> Based on the garage payment data, these garages are never full with maximum occupancy rates of 75%, 72%, and 38% for Marina, Fillmore and Mission, respectively.

Consumer  $i$  may decide to stop searching for on-street parking ( $a = 2$ ). She then faces two options: park at the garage or choose the outside option (e.g., give up the trip completely, park at a private garage, etc.). She chooses the option that maximizes her utility: <sup>10</sup>

$$u_i(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}; 2) = \max(V_i^{\text{garage}}, V_i^{\text{o}}) \equiv V_i^{\text{garage|o}}.$$

Alternatively, consumer  $i$  may decide to park on-street ( $a = 1$ ). Then, if the block is available, i.e.,  $A_{rtdb_k} = 1$ , she gets:  $u_i(b_k, P_{rtdb_k}, 1, \epsilon_{irtdb_k}; 1) = v_{irt} + \epsilon_{is} - \eta_i d(b_i^*, b_k) - \theta_i P_{rtdb_k} h_i \equiv V_i^{\text{park}}(b_k, P_{rtdb_k})$ . If there is no parking spot available in the block, i.e.,  $A_{rtdb_k} = 0$ , then she cannot park there and we denote her utility from parking by negative infinity:  $u_i(b_k, P_{rtdb_k}, 0, \epsilon_{irtdb_k}; 0) = -\infty$ .

Finally, if consumer  $i$  decides to continue to search ( $a = 0$ ), she gets the expected utility:

$$\begin{aligned} & u_i(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}; 0) \\ &= -s_i \epsilon_{irtdb_k} + E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) \mid (b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \right], \\ &\equiv -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}), \end{aligned}$$

where  $s_i$  is consumer  $i$ 's per block search cost and  $\epsilon_{irtdb_k}$  is the shock associated with the search cost  $s_i$ . That is, the actual cost incurred if consumer  $i$  continues to search,  $s_i \epsilon_{irtdb_k}$ , depends on both the per block search cost,  $s_i$ , which is known to the consumer before searching, and the search cost shock,  $\epsilon_{irtdb_k}$ , which is only realized after she arrives at block  $b_k$ . The shocks,  $\epsilon_{irtdb_k}$ , are i.i.d across consumers, regions, time, day and blocks, and follow a standard log-normal distribution (i.e.,  $\log(\epsilon_{irtdb_k})$  follows the standard normal distribution).<sup>11</sup> The log-normal distribution guarantees that the overall search cost  $s_i \epsilon_{irtdb_k}$  is non-negative (a customer would not give up an available parking spot because she suddenly “enjoys” searching). The expectation denotes the expected utility from continuing to search once a customer arrives at block  $b_k$ . It is taken with respect to the conditional distribution of state variables in the next period ( $b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}$ ) given the current state variables ( $b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}$ ).

A consumer that continues to search follows a random walk strategy (with no immediate return) and will arrive at one of the adjacent blocks randomly. Ideally, shocks would be specific to each

<sup>10</sup> We assume that the set of outside options available in the second stage once the search starts is the same as in the first stage. Ideally, we would allow them to differ, as once the search starts, some outside options could be less appealing. However, we cannot empirically distinguish between customers who choose the outside option in the first stage and those who do so in the second stage, and therefore we assume that the two are the same. It is also possible to assume that once the search starts, the outside options are no longer available. Our conclusions are robust to this alternative specification.

<sup>11</sup> As we discuss in the estimation section,  $s_i$  also follows a log-normal distribution. Therefore,  $s_i \epsilon_{irtdb_k}$  is log-normally distributed. Note that without observing individual search paths, it is impossible to estimate the variances of  $s_i$  and  $\epsilon_{irtdb_k}$  separately. We therefore standardize the distribution of  $\epsilon_{irtdb_k}$  to a standard log-normal distribution.

of the adjacent blocks. Although intellectually appealing, block-specific shocks introduce three unobserved state variables (most blocks have three adjacent blocks at each end) and increase the estimation complexity substantially. Without much loss of generality, we introduce one shock to the search cost, which captures, for example, the general traffic condition that makes a consumer more or less willing to continue to search. To introduce randomness to the search path itself, a customer who continues to search elects one of the adjacent blocks randomly. The random walk model allows for idiosyncratic shocks that affect a consumer's stopping decision without over-complicating the estimation.<sup>12</sup>

### *Evolution of States and Consumer Beliefs*

As we discussed, the evolution of the state variable  $b_k$  follows a random walk with no immediate return. The direction of driving at the initial block is generated randomly and there is an equal probability to transition from the current block to any of the adjacent blocks.<sup>13</sup> Let  $B_{rb_k}$  be the set of adjacent blocks accessible from the current block  $b_k$ , and  $|B_{rb_k}|$  be the number of adjacent blocks. The joint evolution of state variables  $P_{rtdb_k}$  and  $A_{rtdb_k}$  depends on the region, time and day, and the location of the current block. For the evolution of the search cost shock, recall that  $\epsilon_{irtdb_k}$  is i.i.d. across consumers, regions, times, days, and blocks and is independent from  $P_{rtdb_k}$  and  $A_{rtdb_k}$ . Therefore, if a consumer decides to continue to search (i.e.,  $a = 0$ ), the transition probability is:

$$Pr(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; 0 | b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \\ = \begin{cases} \frac{1}{|B_{rb_k}|} f_{rtdb}^{P,A}(P_{rtdb_{k+1}}, A_{rtdb_{k+1}} | b_k, P_{rtdb_k}, A_{rtdb_k}) f^\epsilon(\epsilon_{irtdb_{k+1}}), & \text{if } b \in B_{rb_k} \\ 0, & \text{otherwise,} \end{cases}$$

where  $f_{rtdb}^{P,A}$  and  $f^\epsilon$  are the density functions of the state variables.

Consumers form rational expectations about the price, availability and search cost shock distributions. A fully rational consumer would need to possess an extreme level of sophistication: not only would she form rational expectations of the availability, price and search cost shock at each specific block, she would also form rational expectations of the spatial correlations of these state variables, and would therefore update her belief about the distribution of these state variables at a future block based on the observed states of all previous blocks visited. This level of rationality

<sup>12</sup> Although the random walk model abstracts away from some decisions (e.g., which block to drive to based on different beliefs about the expected utility derived from each adjacent block), it does offer reasonable levels of complexity and nuance. Given the many transient random factors at play (traffic, blockage, road condition, traffic lights, emotions, etc.) and limited deliberation time in traffic, it is possible that a random walk may actually be a more realistic model of consumer behavior.

<sup>13</sup> For simplicity of exposition, the notation ignores the direction of driving (i.e., which block a customer searches next), but we simulate driving directions in the estimation.

is undesirable: not only does it introduce a substantial computational burden to the estimation of a dynamic model with multiple state variables, but this level of rationality also expects too much from a consumer. Instead of assuming full rationality, we simplify beliefs and decisions by letting consumers' beliefs of  $P_{rtdb_k}$  and  $A_{rtdb_k}$  be i.i.d. across block  $b_k$ . That is, consumers still have different expectations about the price and availability in different hours of a day, on different days, and in different regions. However, within a region and at a given time  $t$  and day  $d$ , all blocks appear *ex-ante* the same. To explain, taking availability as an example, a consumer forms an expectation that all blocks have the same probability to be available and that this expectation is consistent—it equals the observed average probability of a block being available across all blocks in the region at time  $t$  on day  $d$ . Specifically,  $\phi_{rtdb}$  is the probability that block  $b$  is available (i.e., at least one spot is empty) in region  $r$  at time  $t$  on day  $d$ , and  $\tilde{\phi}_{rtdb}$  is consumers' belief of availability. Then  $\tilde{\phi}_{rtdb} = \frac{\sum_{b \in B_r} \phi_{rtdb}}{|B_r|} \equiv \phi_{rtd}$ . Of course, whether a consumer finds a block available is based on the real-time availability of the block, rather than on average availability. In other words, consumers' beliefs are correct on average, but not for a particular instance. As with realized availability, consumers form rational expectations with respect to prices. Consumers do not know whether the nearby blocks are priced lower than the current block they are at, and only learn how much they will be paying once they arrive at a specific block. In fact, as we show later, the potential rate difference may motivate them to search for better prices in new blocks.

Availability  $\phi_{rtdb}$  is not directly observable but we can derive it from block-level utilization by modeling the block as an  $M/G/s_{rb}/0$  loss system:  $s_{rb}$  is the number of parking spaces at block  $b$ , the number 0 indicates that the maximum queue length is zero (it is a loss system), so that consumers who find that the block is full do not wait in the block for a spot to become available, the arrival rate to block  $b$  at time  $t$  on day  $d$  is  $\lambda_{rtdb}$  and the mean parking time is  $1/\mu_{rtdb}$ . The arrival rate follows a Poisson process but the time spent parking can follow any distribution. From the Erlang loss formula, the probability that a consumer can successfully park, i.e., she is not “lost”, is:

$$\phi_{rtdb} = 1 - \frac{\left(\frac{\lambda_{rtdb}}{\mu_{rtdb}}\right)^{s_{rb}}}{s_{rb}!} \bigg/ \sum_{k=0}^{s_{rb}} \frac{\left(\frac{\lambda_{rtdb}}{\mu_{rtdb}}\right)^k}{k!}. \quad (1)$$

Since only a fraction  $\phi_{rtdb}$  of arriving consumers can be served, utilization is:

$$\varphi_{rtdb} = \frac{\phi_{rtdb} \lambda_{rtdb}}{s_{rb} \mu_{rtdb}}. \quad (2)$$

Rearrange and substitute Equation (1) in Equation (2),  $\phi_{rtdb}$  is implicitly defined by:

$$\phi_{rtdb} = 1 - \frac{\left(\frac{\varphi_{rtdb} s_{rb}}{\phi_{rtdb}}\right)^{s_{rb}}}{s_{rb}!} \bigg/ \sum_{k=0}^{s_{rb}} \frac{\left(\frac{\varphi_{rtdb} s_{rb}}{\phi_{rtdb}}\right)^k}{k!} \quad (3)$$

The expected value that a customer who continues to search derives is:

$$\begin{aligned}
& V_i^{\text{search}}(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \\
&= E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) \mid (b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \right] \\
&= E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) \mid b_k \right] \equiv V_i^{\text{search}}(b_k).
\end{aligned}$$

To explain, the expected value from search,  $V_i^{\text{search}}(b_k)$ , is a function of the current block,  $b_k$ , and it varies across consumers due to differences in parameters and ideal locations. The second equality holds because consumers do not update their beliefs of price and availability, and because the search shocks,  $\epsilon_{irtdb_k}$ , are i.i.d across blocks.

To ensure the robustness of the model and results, we relax the assumptions that consumers have identical beliefs across blocks and do not update them. First, we relax the assumption that consumers have identical price and availability beliefs for all blocks. We analyze an alternative partial-knowledge model, in which we group blocks into three levels (high, medium, and low) according to their popularity before the implementation of congestion pricing. Thus, consumers know the popularity level of a block, and form respective conditional rational expectations of all state variables. Second, we relax the assumption that consumers do not possess the sophistication to update their beliefs. We estimate a variant of the model in which we allow consumers to update their beliefs *once* after they arrive to their ideal location and observe whether it is available or not. We find that the results of our main model are consistent with these modelling specifications. Details of the two alternative models and results are in Appendices C.2 and C.3, respectively.

### *Optimal Decision Rule*

Consumer  $i$ 's optimal decision rule,  $a_i^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k})$ , can be characterized as follows:

- If the current block is available, i.e.,  $A_{rtdb_k} = 1$ , a consumer can choose from three potential actions: continue to search ( $a = 0$ ), park at the current location ( $a = 1$ ), or abandon on-street parking ( $a = 2$ ). A consumer chooses the action that gives her the highest utility:

$$a_i^*(b_k, P_{rtdb_k}, 1, \epsilon_{irtdb_k}) = \begin{cases} 0, & \text{if } -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k) > \max(V_i^{\text{garage}|\text{o}}, V_i^{\text{park}}(b_k, P_{rtdb_k})) \\ 1, & \text{if } V_i^{\text{park}}(b_k, P_{rtdb_k}) \geq \max(V_i^{\text{garage}|\text{o}}, -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k)) \\ 2, & \text{otherwise.} \end{cases}$$

- If the current block is unavailable, i.e.,  $A_{rtdb_k} = 0$ , a consumer has two options to choose from: continue to search ( $a = 0$ ) and stop searching ( $a = 2$ ).

$$a_i^*(b_k, P_{rtdb_k}, 0, \epsilon_{irtdb_k}) = \begin{cases} 0, & \text{if } -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k) > V_i^{\text{garage}|\circ} \\ 2, & \text{otherwise.} \end{cases}$$

Let  $u_i^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) = \max_{a=\{0,1,2\}} u_i(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}; a)$  be the maximum utility a consumer could get after she arrives at block  $b_k$ . This utility depends on whether the block is available and can be computed recursively. Derivations of the utilities are in Appendix A.

## 4.2. Choice of Driving

Finally, we discuss consumer  $i$ 's initial decision. In the first stage consumer  $i$  chooses among three options: (1) drive to the region and search for on-street parking, (2) drive to the region but park directly at a public garage, and (3) choose the outside option. We have specified the utilities from the last two options in the previous section. It remains to derive the expected utility from the first option. The expected utility of consumer  $i$  who decides to drive and search for on-street parking is  $u_{is} = \max_{b_1 \in B_r} E_{(P_{rtdb_1}, A_{rtdb_1}, \epsilon_{irtdb_1})} u_i^*(b_1, P_{rtdb_1}, A_{rtdb_1}, \epsilon_{irtdb_1})$ , where  $b_1$  is the block from which a consumer starts to search ( $k = 1$ ). Consumer  $i$  chooses an initial block  $b_1$  from the set of blocks in the region,  $B_r$ , to maximize her expected utility-to-go. The expectation is taken over the state variables  $(P_{rtdb_1}, A_{rtdb_1}, \epsilon_{irtdb_1})$ , because they are unknown at the time she makes the decision—consumer  $i$  only observes the realizations of these random variables after she arrives at block  $b_1$ .

In the first stage, consumer  $i$  chooses the option that brings her the highest expected utility among  $u_{is}$ ,  $u_{ig}$ , and  $u_{i0}$ , where

$$\begin{aligned} u_{is} &= \max_{b_1 \in B_r} E_{(P_{rtdb_1}, A_{rtdb_1}, \epsilon_{irtdb_1})} u_i^*(b_1, P_{rtdb_1}, A_{rtdb_1}, \epsilon_{irtdb_1}), \\ u_{ig} &= v_{irtd} + \epsilon_{ig} - \eta_i d(b_i^*, b_g) - \theta_i P_{bg} h_i, \\ u_{i0} &= \epsilon_{i0}. \end{aligned}$$

Because we assume identical and rational beliefs across blocks, all blocks are *ex-ante* the same except for their distance to the ideal location. Hence, consumer  $i$ 's best starting location is her ideal location  $b_i^*$ . In a version of the model in which beliefs are stratified across blocks, consumers may not always choose their ideal locations to start their search from. For example, a customer may want to start at a less congested block located far from her ideal location, if she has a high search cost and a low distance disutility cost. (See Appendix C.2 for details.)

## 5. Identification and Estimation

### 5.1. Destination Model

Prior to discussing how we identify and estimate the model parameters, we describe the specification of the destination distribution. We specify the distribution of consumers' destinations as functions of block-level business densities in each region. Since we have annual business density data, we allow the distribution of consumers' destinations to vary by year; the subscript  $y$  accounts for potential changes in business density across years.<sup>14</sup> The fraction of consumers whose destination is block  $b$  in year  $y$ ,  $\omega_{ry}(b)$ , is defined as:

$$\omega_{ry}(b) = \frac{\exp\left(\kappa_r^0 + \kappa_r^A \log N_{ry}^A(b) + \kappa_r^B \log N_{ry}^B(b) + \kappa_r^C \log N_{ry}^C(b) + \kappa_r^{D+} \log N_{ry}^{D+}(b)\right)}{\sum_{b=1}^B \exp\left(\kappa_r^0 + \kappa_r^A \log N_{ry}^A(b) + \kappa_r^B \log N_{ry}^B(b) + \kappa_r^C \log N_{ry}^C(b) + \kappa_r^{D+} \log N_{ry}^{D+}(b)\right)},$$

where  $\log N_{ry}^A$ ,  $\log N_{ry}^B$ ,  $\log N_{ry}^C$  and  $\log N_{ry}^{D+}$  are vectors of the log scaled numbers of businesses of size A (1-4 employees), B (5-9 employees), C (10-19 employees), and D+ (20 employees and above) for region  $r$  in year  $y$ , respectively. This specification guarantees  $0 \leq \omega_{ry}(b) \leq 1$ , and  $\sum_{b=1}^B \omega_{ry}(b) = 1$ . We normalize  $\kappa^0$  to 0 without loss of generality. The parameters left to be estimated are  $\kappa_r^A$ ,  $\kappa_r^B$ ,  $\kappa_r^C$ , and  $\kappa_r^{D+}$  for each region.

The establishments of the businesses are largely exogenous, as they are primarily determined by factors such as availability of commercial real estates, zoning restrictions, licensing requirements, etc. Therefore, they should be little affected by the change of the parking policy or the anticipation of it. Moreover, although changes in destination popularity may lead to business openings or closures, these changes are slow in nature. Indeed we find minimal changes in the number of businesses from the year 2011 to 2012; the correlation of the number of businesses across the two years is over 94% for all regions under study, on par with other regions in the city for which the parking policy did not change.

### 5.2. Identification

The identification of the destination model is straightforward. The distribution of ideal locations,  $\omega$ , is a function of business densities. The exogenous variations in business densities across blocks identify the parameters in the destination model (i.e.,  $\kappa^A$ ,  $\kappa^B$ ,  $\kappa^C$  and  $\kappa^{D+}$ ).

We identify consumer attributes by exploiting parking patterns in the fixed-pricing and congestion pricing periods. We assume that the underlying ideal location distribution does not change in the same region, time and day. To eliminate seasonality effects, we use the same months during the

<sup>14</sup> As a robustness check, we allow the coefficients  $\kappa$  to vary by months, time of day (i.e., morning, noon, or afternoon), weekday vs. weekend by including interactions of these variables and the numbers of businesses. We find that the estimated coefficients of the interaction terms are not statistically significant.

before and after periods. The three key consumer attributes that we are interested in are search cost, distance disutility, and price sensitivity. Similar to Hortacsu and Syverson (2004), Hong and Shum (2006), and Kim et al. (2010), even though we do not directly observe consumers' search paths, we are able to identify search costs and other relevant parameters.<sup>15</sup> We explain which variations in our data drive the identification of each parameter:

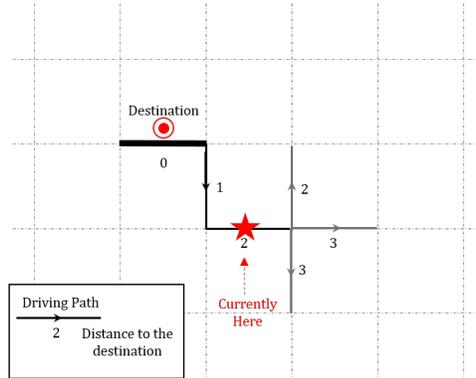
*Separation of Price Sensitivity from Search Cost and Distance Disutility.* Price sensitivity is identified through the exogenous variation of prices from the SFpark program and the resulting variations in parking locations. To illustrate, consider a simple case with one block and one garage in the region. Naturally, in this case the block is the ideal location for all consumers, but some will have to park at the garage on a congested day. What differentiates on-street from off-street parking is the parking fee and the distance disutility. Although the distance disutility does not change before and after the implementation of congestion pricing, the parking fee does. When the on-street price increases following the implementation of congestion pricing, more consumers park at the garage and fewer park on-street, and vice versa. The extent to which price changes can induce the reallocation of parking between on- and off-street identifies price sensitivity (relative to distance disutility).

*Separation of Search Cost from Distance Disutility.* There are two sources of variation that separate search cost and distance disutility. The first is the extent to which demand shifts to the garage rather than to nearby, less congested blocks. By choosing to park at the garage, a consumer avoids an additional search cost but usually incurs a greater walking distance to her destination. By choosing to continue to search, a consumer incurs the search cost but may reduce the walking distance, if she finds a parking space nearby. Therefore, if we observe that parking demand shifts to the garage rather than to nearby, less congested blocks, we can infer that consumers are relatively more sensitive to the inconvenience induced by search than to walking.

The second source of variation stems from the imperfect correlation between the change in search cost and the respective change in walking distance. If a consumer chooses to continue to search, the total number of blocks searched always increases by one regardless of which nearby block she visits next. However, the walking distance between the next block and her destination may increase or

<sup>15</sup> These studies estimate search related parameters in dynamic models without observing the actual search paths. Hortacsu and Syverson (2004) model how investors search over funds with varying attributes, and estimate heterogeneous search costs using market share data and price data only. Hong and Shum (2006) develop a method to uncover heterogeneous search costs with price data alone, using equilibrium conditions from sequential and non-sequential search models. Kim et al. (2010) estimate consumer search costs in online retailing with view-rank data. Though the exact models and the variations used to identify search parameters differ, as Hortacsu and Syverson (2004) write, these papers demonstrate "how aggregate data can be used to identify and estimate search costs separately from product differentiation, with particular attention to minimizing the impact of functional form restrictions." In fact, our data is more detailed than in those papers, because we observe each decision maker's final decision and price variations over time due to congestion pricing, not just aggregate market shares and prices.

decrease depending on the path she takes (see Figure 5).<sup>16</sup> This seemingly subtle variation comes from the fact that the search is conducted on a two-dimensional space. If instead the search was conducted on a unidimensional line, then the number of blocks searched would perfectly correlate with the distance from the ideal location. In this case, it would be more difficult to determine whether a redistribution of demand is caused by aversion to search or to walking.



**Figure 5 Identification Illustration: Separate Search Cost and Distance Disutility**

### 5.3. Moment Conditions

The primitives of the model that we wish to estimate are: coefficients of the destination model,  $\kappa$ , trip valuation parameters,  $\alpha$  and  $\beta$ , the *joint* distribution of search cost, distance disutility and price sensitivity,  $(s_i, \eta_i, \theta_i)$ , and the standard deviation of the utility shocks,  $\sigma$ . We assume that the joint distribution of  $(s_i, \eta_i, \theta_i)$  follows a multivariate lognormal distribution  $\ln N(\mu_{s,\eta,\theta}, W_{s,\eta,\theta})$ , where  $\mu$  and  $W$  are the mean and the variance-covariance matrix of the corresponding normal distribution. All parameters are region specific. We jointly estimate ideal location distributions and the consumer attribute parameters using data from both the fixed-pricing and congestion-pricing periods. We then use the ideal location distributions and the consumer attributes that we estimated as inputs for the counterfactual analyses.

By estimating the joint distribution, we allow a consumer's search cost, distance disutility and price sensitivity to be correlated, and we estimate the correlations empirically. Let  $\Theta = (\mu_{s,\eta,\theta}, W_{s,\eta,\theta}, \alpha, \beta, \sigma, \kappa)$ . Because the scale of utility is irrelevant to choices, not all parameters can be identified. Consumers compare relative values among options, so the exact scale of utility is irrelevant, and the parameters are estimated relative to a parameter that is normalized to 1. To measure welfare in dollars, we choose to normalize the mean price sensitivity to 1.

<sup>16</sup> In Figure 5, a consumer may turn left, right, or continue straight from the current block. If she turns left, it doesn't change her distance from her ideal location; if she turns right or continues straight, she will be one block farther from her ideal location. That is, with a random walk, the expected increase in distance to ideal location if she continues to search is  $(\frac{1}{3} \times 2 + \frac{1}{3} \times 3 + \frac{1}{3} \times 3) - 2 = \frac{2}{3}$ , while the expected increase in the number of blocks searched is exactly 1.

We use Simulated Method of Moments (SMM) to estimate  $\Theta$ . First proposed by McFadden (1989), SMM is conceptually identical to the more commonly used GMM, except that with SMM the moments are calculated using model-based simulations rather than directly from the model. Researchers use SMM instead of GMM when it is difficult (or impossible) to derive moment conditions from the model directly. For example, in our setting it is impossible to derive an explicit analytical expression for the equilibrium. Therefore, we solve it numerically using simulations. Given an admissible set of parameters  $\Theta$ , we simulate the driving decision and parking location for every consumer in the market. We then calculate the number of people who park at each block and in the garage, and customers' total parking duration, and match these to the corresponding moments we observe in the data. Specifically, we calculate the differences between the observed and the simulated outcomes for the fixed- and congestion-pricing periods, given exogenous variables  $V$  and parameters  $\Theta$ , as:

$$g(V; \Theta) = \begin{bmatrix} q_{rtd}^{\text{fixed}}(b) - \tilde{q}_{rtd}^{\text{fixed}}(b; \Theta) \\ Q_{rtd}^{\text{fixed}}(b) - \tilde{Q}_{rtd}^{\text{fixed}}(b; \Theta) \\ q_{rtd}^{\text{dynamic}}(b) - \tilde{q}_{rtd}^{\text{dynamic}}(b; \Theta) \\ Q_{rtd}^{\text{dynamic}}(b) - \tilde{Q}_{rtd}^{\text{dynamic}}(b; \Theta) \end{bmatrix}, \forall t, d, b = \{1, 2, \dots, |B_r|, g\},$$

where  $q$  is the number of people who park,  $Q$  is the total minutes parked at each block  $b$  or in the garage  $g$  ( $Q$  and  $q$  are observed outcomes and  $\tilde{Q}$  and  $\tilde{q}$  are simulated outcomes) and  $V$  is the matrix of indicators for each block/the garage, month (May, June, and July), weekend, and hour (morning, noon, and afternoon). We estimate the parameters  $\hat{\Theta}$  based on these orthogonality conditions:

$$\mathbb{E}[V'g(V; \Theta)] = \mathbf{0}. \quad (4)$$

We have  $(|B| + 1) \times 4 \times 7$  moment conditions overall. (See Appendix B for details on the estimation procedure.)

## 6. Estimation Results, Welfare Analysis, and Robustness Tests

In this section, we report the estimated availability, ideal location distributions and model parameters. Using the estimates, we calculate welfare changes from before and after the implementation of congestion pricing. Finally, we perform multiple sets of robustness tests to ensure that our results are driven by data variations rather than by specific modeling assumptions.

### 6.1. Estimation Results

We calculate real-time availability for each hour and each block in our sample periods using occupancy data obtained from the parking sensors. Table 2 summarizes the availability estimates for

**Table 2 Summary of Availability Estimates**

	Marina					Fillmore					Mission				
	Total	High	Mid	Low	Garage	Total	High	Mid	Low	Garage	Total	High	Mid	Low	Garage
Availability—Before	0.92 (0.06)	0.88 (0.06)	0.93 (0.05)	0.98 (0.03)	1.00 (0.00)	0.92 (0.07)	0.86 (0.09)	0.93 (0.06)	0.96 (0.04)	1.00 (0.00)	0.99 (0.02)	(-) (-)	0.99 (0.02)	0.99 (0.06)	1.00 (0.00)
Availability—After	0.93 (0.06)	0.89 (0.07)	0.93 (0.05)	0.98 (0.02)	1.00 (0.00)	0.92 (0.08)	0.87 (0.09)	0.93 (0.07)	0.96 (0.04)	1.00 (0.00)	0.98 (0.03)	(-) (-)	0.98 (0.02)	0.97 (0.03)	1.00 (0.00)

\* Standard deviations are in parentheses. High, medium and low utilization blocks are defined using average occupancy rate greater than 80%, between 60% and 80%, and below 60%, respectively. “Before” refers to our sample period before congestion pricing: April to July in year 2011, “after” refers to our sample period after congestion pricing: April to July in year 2012.

high-, medium-, and low-utilization blocks in each region during the before and after periods. The estimates show similar patterns as the occupancy rates in Table 1.

We also estimate the per-block price sensitivity, search cost, and distance disutility, their covariance matrix, the four parameters in the destination model, and the seven parameters that affect trip valuation. Table 3 presents the results. To interpret the magnitudes of the estimated coefficients, we convert the estimates of search cost and distance disutility to dollar values. In the main specification, for example, we estimate that it costs a median consumer in Marina approximately \$3.96 to search an additional block and \$3.10 to park one block away from their destination (all measured by first and third quartiles). We obtain similar estimates for the other two regions. We also examine the estimated correlations between search cost, distance disutility and price sensitivity. As expected, less price sensitive consumers also value their time more (there are negative correlations between price sensitivity and search costs). Moreover, customers who dislike search also dislike parking farther away from their destinations (there is a positive correlation between search cost and distance disutility).

Based on the estimated model, we calculate the average price elasticity (i.e., percentage change in a block’s occupancy rate as a result of one percent change in its price) to be -0.36, -0.24 and -0.35 for Marina, Fillmore and Mission, respectively. Interestingly, the estimates are of the same magnitude as those reported in Pierce and Shoup (2013) and Ottosson et al. (2013) (from -0.8 to -0.4). To evaluate model fit, we compare predicted and observed moments for each block. Figure 9 demonstrates a close moment fit by blocks. We also calculate the amount of variation in the observed moments that can be explained by the model. At the hourly level, the model explains an average of between 22% and 40% of the variations in the data. Lastly, we conduct in-sample and out-of-sample analyses by randomly selecting two-thirds of time and day in each region to be used for the in-sample analysis and the remaining one third for out-of-sample analysis. The in-sample R-squares are 42.15%, 40.65%, 21.17% for Marina, Fillmore and Mission, respectively, while the out-of-sample R-squares are 40.63%, 39.03%, and 21.79% for each region, respectively.

**Table 3 Model Estimates**

	Marina	Fillmore	Mission
Search cost (log-scaled) mean	1.38 (0.11)	0.78 (0.10)	0.98 (0.09)
Search cost (log-scaled) sd	1.22 (0.13)	1.07 (0.13)	1.87 (0.06)
Distance disutility (log-scaled) mean	1.14 (0.10)	2.07 (0.08)	1.72 (0.07)
Distance disutility (log-scaled) sd	0.05 (0.03)	0.05 (0.02)	0.23 (0.03)
Price sensitivity (log-scaled) mean	Normalize to 0		
Price sensitivity (log-scaled) sd	0.90 (0.10)	0.29 (0.13)	0.52 (0.07)
$\epsilon$ sd	0.90 (0.10)	0.74 (0.21)	0.11 (0.21)
Search cost $\times$ distance disutility corr	0.00 (0.01)	-0.02 (0.03)	0.07 (0.02)
Search cost $\times$ price sensitivity corr	-1.51 (0.10)	-1.86 (0.09)	-1.81 (0.08)
Distance disutility $\times$ price sensitivity corr	0.60 (0.09)	0.60 (0.05)	0.46 (0.07)
Trip valuation— $\alpha$	4.51 (0.39)	3.17 (0.28)	3.61 (0.24)
Trip valuation—intercept	9.74 (0.63)	6.37 (0.41)	7.15 (0.32)
Trip valuation—May	Baseline		
Trip valuation—June	0.10 (0.16)	-0.38 (0.09)	0.24 (0.09)
Trip valuation—July	0.13 (0.16)	-0.02 (0.09)	0.30 (0.10)
Trip valuation—Weekday	Baseline		
Trip valuation—Weekend	-0.31 (0.15)	2.04 (0.02)	1.37 (0.13)
Trip valuation—morning	Baseline		
Trip valuation—noon	0.06 (0.16)	0.58 (0.10)	0.59 (0.10)
Trip valuation—afternoon	0.39 (0.17)	0.96 (0.13)	0.62 (0.10)
Destination Model— $\kappa^A$	-0.34 (0.03)	-0.11 (0.01)	0.20 (0.02)
Destination Model— $\kappa^B$	0.56 (0.04)	0.47 (0.02)	-0.01 (0.02)
Destination Model— $\kappa^C$	-0.34 (0.05)	0.23 (0.02)	0.04 (0.02)
Destination Model— $\kappa^{D+}$	0.21 (0.04)	-0.05 (0.31)	-0.18 (0.02)
Search cost dollar value	\$3.96	\$2.19	\$2.67
Distance disutility dollar value	\$3.10	\$7.98	\$5.60

\* The dollar value interval displays the median of the distribution over commuters.

## 6.2. Welfare and Search Externality

*Welfare.* Using the model estimates, we quantify the effect of congestion pricing on both consumer and social welfare. We denote the actual utility that consumer  $i$  obtains by  $u_i^{\text{actl}}$ , the actual parking location for consumer  $i$  by:

$$b_i^{\text{actl}} = \begin{cases} b, & \text{if consumer } i \text{ parks at block } b \text{ eventually, } b \in B_r; \\ b_g, & \text{if consumer } i \text{ parks at the garage eventually;} \\ o, & \text{if consumer } i \text{ chooses the outside option eventually;} \end{cases}$$

and the actual number of searches that consumer  $i$  has made by  $N_i$ . Then,

$$u_i^{\text{actl}} = \begin{cases} v_{irt} + \epsilon_{is} - N_i \left( \sum_{k=1}^{N_i} s_i \epsilon_{irt} b_k \right) - \eta_i d(b_i^*, b_i^{\text{actl}}) - \theta_i P_{rt} b_i^{\text{actl}} h_i, & \text{if } b_i^{\text{actl}} = b, b \in B_r; \\ u_{ig} - N_i \left( \sum_{k=1}^{N_i} s_i \epsilon_{irt} b_k \right), & \text{if } b_i^{\text{actl}} = b_g; \\ u_{i0} - N_i \left( \sum_{k=1}^{N_i} s_i \epsilon_{irt} b_k \right), & \text{if } b_i^{\text{actl}} = o. \end{cases}$$

We calculate consumer surplus,  $CS$ , and social welfare,  $SW$ , as:  $CS = \sum_{i=1}^{M_r} \frac{1}{\theta_i} u_i^{\text{actl}}$  and  $SW = CS + \sum_{i=1}^{M_r} P_{rt} b_i^{\text{actl}} h_i$ . Following the literature (see Meijer and Rouwendal 2006 and references therein), we divide utility by price sensitivity,  $\theta_i$ , such that welfare is expressed in dollar values. As is customary, in calculating social welfare, we view parking payment as a transfer from consumers to the government, which is then distributed back to the local community in various ways.

We report welfare and search traffic in Table 4. Note first that while customers search for an available parking spot irrespective of the pricing strategy, congestion pricing introduces another type of search—search for a better price. That is, due to the lack of complete information about prices, customers who find an available spot at a higher than expected price, may forgo that space and continue searching thereby increasing costs (see the lower panel of Table 4). While we find that in the regions we examined the cost to search for an available parking spot outweighs the cost to search for a better price, it is possible that in cities or regions with more price sensitive customers, prices will be more dispersed and the search for prices may be more pervasive, increasing price-based search cost. Policy makers should therefore account for this cost when considering a change in pricing strategy.

In the main specification, following the implementation of congestion pricing, consumer surplus increased by \$41.05 and \$19.50 per a hundred consumers, an equivalent of 55.39% and 25.44% of total payment, in Marina and Fillmore, respectively. However, consumer surplus decreased by \$31.76 per a hundred consumers, an equivalent of 45.02% of total payment, in Mission. We observe the same directional changes in social welfare.

**Table 4 Welfare and Search Externality**

	Marina		Fillmore		Mission	
	Before (Uniform Pricing)	After ( <i>SFpark</i> Pricing)	Before (Uniform Pricing)	After ( <i>SFpark</i> Pricing)	Before (Uniform Pricing)	After ( <i>SFpark</i> Pricing)
Panel A: Welfare Calculations						
Distance Disutility	84.42	71.94	527.45	519.69	47.43	58.36
Search Cost	71.94	47.41	156.41	139.60	2.87	8.74
Payment	74.12	91.92	76.64	85.90	70.54	75.11
Trip Valuation	689.09	710.93	1178.60	1182.79	692.36	681.97
Consumer Surplus Change		41.05		19.50		-31.76
Social Welfare Change		58.86		28.76		-27.18
Panel B: Summary Statistics of Consumer Actions and Search Traffic						
Go with the Car (%)	50.19	48.22	51.62	50.10	50.44	48.61
Park at Garage (%)	4.95	5.63	14.07	14.55	2.84	3.66
Search Availability (%)	13.23	8.04	14.84	13.03	0.51	1.40
Search Price (%)	0.00	0.18	0.00	0.18	0.00	0.34
Total # of Searches	13.23	8.22	14.84	13.21	0.51	1.74

Notes. Consumer surplus is normalized to dollar value at the size of a hundred consumers. Results are based on 50 rounds of simulations.

Where do the differences in welfare originate from? In Marina and Fillmore, following the implementation of congestion pricing, consumers incur much lower search costs and lower distance disutility which offset the increase in payments. The same does not hold in the Mission district. There, the implementation of congestion pricing increased not only total payment, but also search costs and distance disutility. Several effects may contribute to the increase in search costs and distance disutility. First, congestion pricing does not only decrease congestion of highly popular blocks, it also increases congestion at low utilization blocks—high prices in popular blocks lead more customers to less congested blocks in search for a better price, but this results in increased congestion in these blocks. The increase in congestion may increase both search costs—search for a better price and search for availability in less congested regions due to the externality imposed by consumers searching for a better price) and distance disutility. Second, consumers may find it less attractive to drive to their destination if they anticipate higher prices.

The overall effect is that both social welfare and consumer surplus decrease. The inconsistent welfare implications in different regions highlight the critical tradeoff between the desire for high utilization and the aversion for congestion. From the perspective of resource utilization, social planners would like to attract as many customers as possible and keep utilization high. However, high utilization generates congestion, which reduces the utility that each consumer obtains from accessing the resource. We further examine the reasons that congestion pricing leads to lower welfare in Mission as well as pricing policies that may increase welfare in Section 7.1.

Finally, following the implementation of congestion pricing, the total number of searches decreased by 37.8% and 10.96% in Marina and Fillmore, respectively. While we do not have traffic data, these numbers suggest that the program contributes to a reduction in traffic levels. Decreased traffic has implications on gas usage, pollution and accidents, all of which are likely to decrease

as well, contributing to a more sustainable world. In addition to the negative externalities that congestion pricing helps mitigate, it may also result in positive externalities. For example, better parking availability may have implications on the economics of local businesses in these regions.

### 6.3. Heterogeneous Welfare Impacts

To further understand the effects of congestion pricing on welfare, we take advantage of the heterogeneity of consumers and examine changes in welfare by different consumer segments in Marina and Fillmore in Table 5. First, we compare changes in consumer surplus by price sensitivity levels. Consumers with low price sensitivity benefit from congestion pricing more than consumers with high price sensitivity. This makes intuitive sense because consumers who are not very sensitive to price are more willing to trade off monetary payments with congestion costs (search and distance disutility). Second, we compare changes in consumer surplus by the popularity of ideal locations. We define an ideal location's popularity level based on the occupancy rate in year 2011; the popular locations account for more than 64% of consumers in Marina and Fillmore. Consumers whose destinations are at more popular locations are more likely to benefit from congestion pricing. To explain, after the implementation of congestion pricing, popular blocks become more available, which allows consumers who demand these blocks to park there. Therefore, consumers whose ideal locations are popular benefit from congestion pricing more on average. At the same time, with congestion pricing, more consumers park at less popular blocks. As a result, less popular blocks become more congested than with fixed pricing. Consumers whose ideal locations are less popular incur higher search costs and distance disutility and are therefore relatively worse off.

**Table 5** Impacts of Congestion Pricing on Consumer Surplus by Consumer Segment

	Consumer surplus Change (\$)		Fraction of Consumers (%)	
	Marina	Fillmore	Marina	Fillmore
Consumer Segments by Price Sensitivity				
High	-6.40	-11.66	50.00	50.00
Low	47.46	31.16	50.00	50.00
Consumer Segments by Ideal Locations				
High-Popularity Blocks	28.19	24.65	68.52	64.19
Low-Popularity Blocks	12.87	-5.15	31.48	35.81

Notes. Consumer surplus is normalized to dollar value at the size of a hundred consumers. Results are based on 50 rounds of simulations. High (low) price sensitivity is categorized by whether the price sensitivity is greater (less) than the 50th percentile. High and low popularity blocks are defined as the blocks which occupancy rate in year 2011 is greater (less) than the 50th percentile.

### 6.4. Robustness Tests

We conduct multiple robustness tests to ensure that our results are not driven by specific modeling assumptions. We evaluate the robustness of our results along the following dimensions: (1) Market

size. To ensure that our results are not sensitive to the choice of market size, we perform the analyses for several market sizes. (Appendix C.1.) (2) Consumers’ beliefs. Rather than assuming that consumers’ beliefs regarding availability and price are identical across blocks at a given time and day in a region, we allow consumers to form different beliefs depending on the levels of congestion. (Appendix C.2.) (3) Updating beliefs. We allow consumers to update their beliefs about availability if they arrive at the ideal location and find it unavailable. (Appendix C.3.) (4) Parking duration distribution. Due to the irregularities observed in the distribution of parking durations (details explained in Appendix B), we draw parking durations from the observed empirical distribution. The observed distribution is censored. To see whether censoring may affect our conclusions, we simulate different distributions of parking duration based on different censoring levels and introduce additional noise (Appendix C.4). Our conclusions regarding welfare and search traffic changes are robust to all these alternative specifications.

## 7. Counterfactual Analyses

We conduct three counterfactual analyses: We examine what alternative pricing policies may lead to welfare improvement for uncongested regions such as Mission. We examine simpler pricing structures to balance availability-based and price-based searching. Finally, we test how congestion pricing compares to the commonly used policy that limits parking duration, but keeps prices fixed.

### 7.1. Pricing in Uncongested Regions

Recall that the congestion pricing policy implemented by *SFpark* lowered consumer and social welfare in Mission. A critical difference between Mission and the other two regions is that Mission was not very congested even before the implementation. Table 1 shows that even before the implementation of the program, no block in Mission was classified as “high” occupancy. The average occupancy rate was 62% with all blocks reporting below 70% in the before period (April to July in 2011).<sup>17</sup> To compare, before the *SFpark* program was implemented, the average occupancy rate in Marina was 74% and in Fillmore 70%. We also find that occupancy rates are less dispersed geographically in Mission compared to the other regions. This suggests that the primary focus in Mission should not be to reallocate demand across blocks, but to increase block utilization. We therefore hypothesize that lower parking rates may increase welfare in Mission, by increasing the fraction of consumers who drive to the destination.

Specifically, we consider two counterfactuals with lower prices: (1) uniform pricing, where each block is priced equally at a rate which is \$0.50 lower than the average price charged during the

<sup>17</sup> There were blocks with occupancy rates slightly above 80% in other months that year, which led to subsequent price increases.

after period; (2) congestion pricing, in which each block is priced \$0.50 lower compared to the corresponding price during the after period for that block. Lower parking rates introduce a tradeoff: They impact parking costs and entice more consumers to drive. At the same time, the increase in demand leads to more traffic and congestion, which make driving less desirable. In uncongested areas with low utilization, availability may remain high even when utilization increases. Therefore, decreasing prices in such areas could be beneficial if the gain from increased utilization dominates the negative impact of congestion. In overly congested areas, lower rates result in increased congestion without a welfare gain, because availability drops dramatically as utilization increases. We find the new equilibrium following procedures presented in Appendix D, with results shown in Table 6.

**Table 6 Welfare and Search Externality with Lowering Prices Uniformly by \$0.5**

	Before (Uniform Pricing)	After (SFpark Pricing)	Counterfactual I (Uniform pricing \$0.5 lower)	Counterfactual II (SFpark Pricing \$0.5 lower)
<i>Panel A: Welfare Calculations</i>				
Distance Disutility	47.43	58.36	51.67	60.54
Search Cost	2.87	8.74	4.78	9.34
Payment	70.54	75.11	57.16	62.57
Trip Valuation	692.36	681.97	700.76	692.13
Consumer Surplus Change		-31.76	15.63	-11.83
Social Welfare Change		-27.18	2.25	-19.80
<i>Panel B: Summary Statistics of Consumer Actions and Search Traffic</i>				
Go with the Car (%)	50.44	48.61	53.95	52.02
Park at Garage (%)	2.84	3.66	2.91	3.91
Search Availability (%)	0.51	1.40	0.91	1.64
Search Price (%)	0.00	0.34	0.00	0.44
Total # of Searches	0.51	1.74	0.91	2.08

\* Consumer surplus is normalized to dollar value at the size of a hundred commuters. Results are based on 50 rounds of simulations.

We show that lower parking rates increase consumer and social welfare under uniform pricing with lower rates. Much of the gain can be attributed to higher fractions of consumers driving to the destination, i.e., 52–54% as opposed to 49%. Even though consumers incur slightly higher search costs or park farther away, the social gain from the increased total trip valuation more than offsets losses in search costs and distance disutility, confirming our intuition that when the region is underutilized, demand reallocation is secondary to the benefit gained from increased utilization. This result illustrates that it is important to determine whether congestion is a real concern in the region. If it is not, then alternative policies aimed at increasing utilization may lead to more desirable outcomes. With congestion pricing, even when the rates are lower, consumer surplus and social welfare decrease. This is because incomplete information on availability and prices cause consumers to park farther away from their ideal locations and results in higher congestion in some blocks, additional searches, and increased disutility costs.

The results of the counterfactuals combined highlight the importance of treating different types of regions *strategically* differently. Policies that work well in highly congested areas, such as congestion pricing, may not work well in under-utilized regions even if the price levels are lower. These areas lend themselves to fundamentally different policies.

## 7.2. Comparing Simple and Complex Pricing Policies

The complexity of a pricing strategy along with the uncertainty it brings may cause congestion pricing strategies to fail (Bonsall et al. 2007). In our setting, a complex pricing strategy may lead consumers to adopt inefficient search strategies with regard to where to start and where to search. It may also induce search for better prices. To see if a simpler pricing structure may lead to higher welfare, we examine a pricing policy with only three price levels, each corresponding to the high, medium, and low-utilization blocks, respectively. Given that there are only three price levels, it is likely that consumers have perfect information about prices, which is what we assume. To allow for a fair comparison, we set each of the three price levels to equal the average price observed for high-, medium- and low-utilization blocks. We keep the rate constant for the entire study period regardless of the time of the day and day of week.

We solve for the equilibrium under the three-tier pricing policy. The results, presented in Table 7, suggest that the simpler pricing policy achieves higher social welfare and consumer surplus in all three regions, compared to the more complex congestion pricing policy currently in place. Much of this gain can be attributed to higher trip valuation. Moreover, the simpler pricing policy further reduces total search traffic by 0.13 percentage points (or 1.00%) in Fillmore relative to congestion pricing.

Even with the simpler pricing policy, the social and consumer welfare are lower than with fixed pricing in the Mission, which as we argued earlier, illustrates that the primary focus in Mission should not be to reallocate demand across blocks through differentiated pricing, but to increase utilization.

## 7.3. Usage Limits versus Congestion Pricing

The regulation approach (e.g., usage limits or permits) and the market-based approach (e.g., price-based approach) are the two most commonly used approaches in managing public resources. In city parking, most local governments impose parking duration limits to regulate the usage of public parking spaces, but some have recently used congestion pricing to match demand with limited supply. For example, the City of San Francisco previously imposed 2-hour parking limits on most blocks, but relaxed the limit to 4 hours when it decided to pilot the new congestion pricing program in April 2011, i.e., the start of the before period. To compare the two approaches, we examine the

**Table 7 Welfare and Search Externality with Three-Tier Pricing**

	Marina			Fillmore			Mission		
	Before (Uniform Pricing)	After (SFpark Pricing)	Counterfactual (3-Tier Pricing)	Before (Uniform Pricing)	After (SFpark Pricing)	Counterfactual (3-Tier Pricing)	Before (Uniform Pricing)	After (SFpark Pricing)	Counterfactual (3-Tier Pricing)
<i>Panel A: Welfare Calculations</i>									
Distance Disutility	84.42	71.94	83.19	527.45	519.69	520.90	47.43	58.36	58.50
Search Cost	71.94	47.41	69.86	156.41	139.60	176.85	2.87	8.74	8.73
Payment	74.12	91.92	105.51	76.64	85.90	100.85	70.54	75.11	71.50
Trip Valuation	689.09	710.93	778.17	1178.60	1182.79	1237.50	692.36	681.97	683.66
Consumer Surplus Change		41.05	61.01		19.50	20.81		-31.76	-26.60
Social Welfare Change		58.86	92.41		28.76	45.01		-27.18	-25.63
<i>Panel B: Summary Statistics of Consumer Actions and Search Traffic</i>									
Go with the Car (%)	50.19	48.22	48.31	51.62	50.10	49.76	50.44	48.61	49.62
Park at Garage (%)	4.95	5.63	4.82	14.07	14.55	13.75	2.84	3.66	3.32
Search Availability (%)	13.23	8.04	8.66	14.84	13.03	13.08	0.51	1.40	1.48
Search Price (%)	0.00	0.18	0.00	0.00	0.18	0.00	0.00	0.34	0.20
Total # of Searches	13.23	8.22	8.66	14.84	13.21	13.08	0.51	1.74	1.68

\* Consumer surplus is normalized to dollar value at the size of a hundred commuters. Results are based on 50 rounds of simulations.

counterfactual of fixed pricing with a 2-hour parking limit. In this case, if a consumer wants to park for more than the 2-hour limit, she has three options: she could compromise and park for up to 2 hours on street, park at the garage for the entire time demanded, or choose the outside option. A consumer will choose the option that maximize her utility.

Consumers that require long parking durations tend to value the trip more. Limiting their parking duration or blocking them altogether harms welfare by decreasing overall trip valuation. At the same time, usage limits increase availability because they exclude consumers who demand long parking durations, and possibly allow parking spaces to be utilized by more consumers who require short parking durations. Therefore, the overall effect of usage limits on welfare depends on whether the gain from better availability offsets the loss in trip valuation.

Table 8 illustrates that the resulting social welfare with congestion pricing is higher in all regions compared to the social welfare achieved with time limits. The results have an intuitive explanation. To maximize social welfare, a social planner would allocate the parking spaces to consumers who value them the most. Congestion pricing aims at doing exactly that—by charging different prices based on congestion, the policy allocates the more desired spots to high value customers (customers with higher trip valuation and lower search costs, distance disutility and price sensitivity). Prices are only a transfer, so they do not affect social welfare. By contrast, imposing limits on parking durations makes consumers who want to park longer particularly worse off, because they are forced to park for shorter time or seek alternative options that are less desirable. This is especially problematic if consumers' values for the trip are positively correlated with the length of the trip, implying that drivers who value the trip more will likely be hurt the most. Therefore, congestion pricing leads to a more efficient allocation compared to time limits, and we would expect social welfare to be higher with congestion pricing compared to time limits, as we indeed observe.

The comparison of consumer surplus is less intuitive, because the prices charged affect consumer surplus. That introduces a tradeoff: although congestion pricing can allocate demand more efficiently, it does so by charging higher prices on average, and these higher prices (compared to the

**Table 8 Welfare and Search Externality with 2-hour Usage Limit**

	Marina		Fillmore		Mission	
	After (SFpark Pricing)	Counterfactual (2-Hour Time Limit + Uniform Pricing)	After (SFpark Pricing)	Counterfactual (2-Hour Time Limit + Uniform Pricing)	After (SFpark Pricing)	Counterfactual (2-Hour Time Limit + Uniform Pricing)
<i>Panel A: Welfare Calculations</i>						
Distance Disutility	71.94	64.80	519.69	505.29	58.36	33.42
Search Cost	47.41	50.84	139.60	138.59	8.74	1.77
Payment	91.92	66.88	85.90	72.56	75.11	63.74
Trip Valuation	710.93	640.84	1182.79	1159.29	681.97	642.48
Consumer Surplus Change		-41.32		5.26		3.79
Social Welfare Change		-66.37		-8.09		-7.58
<i>Panel B: Summary Statistics of Consumer Actions and Search Traffic</i>						
Go with the Car (%)	48.22	50.43	50.10	51.66	48.61	50.10
Park at Garage (%)	5.63	3.89	14.55	13.60	3.66	2.20
Search Availability (%)	8.04	10.68	13.03	13.39	1.40	0.34
Search Price (%)	0.18	0.00	0.18	0.00	0.34	0.00
Total # of Searches	8.22	10.68	13.21	13.39	0.51	1.74

\* Consumer surplus is normalized to dollar value at the size of a hundred commuters. Results are based on 50 rounds of simulations. Consumer surplus and social welfare changes are evaluated relative to the after period with congestion pricing.

lower uniform price) influence consumer surplus negatively. Which effect dominates depends on many factors such as the time limits imposed, the levels and spread of prices, and market and consumer characteristics. Indeed, we find that the effect on consumer surplus is ambiguous. Relatedly, in a queueing model in which customers choose their time in service, for which parking is a good example, Feldman and Segev (2018) shows that to maximize consumer surplus, a social planner should set the price to zero, but impose time limits. We find a similar result. When comparing congestion pricing with a 2-hour time limit in which parking is free, we find that social welfare is higher with congestion pricing in all regions, but that consumer surplus is higher with time limits but a zero price.

Regardless, both the market-based congestion pricing policy, and the regulation-based approach of limiting parking time, are strategies designed to control congestion, and are likely to decrease utilization. It is imperative to think carefully whether, and if so how to correctly implement them, and the decision depends on the objective that the social planner tried to maximize, as our results suggest.

## 8. Conclusion

Congestion pricing is often considered as an effective tool to match supply with demand. Using data from SFpark, we find evidence that congestion pricing helps to increase parking availability in congested areas, reduces search costs, and allows consumers to park closer to their destinations when they need to. These benefits outweigh the increased payments and lead to an overall increase in consumer surplus and social welfare. Interestingly, we find that congestion pricing can sometimes reduce consumer and social welfare. This happens in areas with relatively low congestion levels, in which improving the overall utilization by setting a proper price level is often more important than reallocating demand by charging variable prices. One solution does not fit all. Therefore, city governments should not apply congestion pricing blindly. Rather, they should diagnose and address the primary concern of each region separately.

Our results also highlight that cities that consider implementing congestion pricing policies should avoid setting unnecessarily complex pricing rules. Even though congestion pricing is intended to reduce search and traffic and increase availability, complicated policies limit consumers' ability to remember and process information, and can lead to inferior decisions. We show that to achieve the best welfare outcome, it is important to balance the desired availability targets with the complexity of the pricing policy—implementing a simpler three-tier pricing policy may be sufficient.

More generally, our learnings from *SFpark* offer important lessons to other public sectors. Both regulation-based and market-based approaches have been used in many public sectors. We provide evidence that the market-based congestion pricing approach generates higher social welfare than the regulation-based usage limit approach, because congestion pricing tends to allocate resources to consumers who value them the most, while usage limits may hurt these consumers more. Although policy decisions are often multi-faceted and it is difficult to account for and measure all possible factors at play, our analysis offers a generalizable methodology and quantifiable results that public sector managers can use to better evaluate the tradeoffs involved.

## References

- Akşin, Z., B. Ata, S. M. Emadi, C-L. Su. 2013. Structural estimation of callers' delay sensitivity in call centers. *Management Science* **59**(12) 2727–2746.
- Anand, K.S., M. F. Paç, S. Veeraraghavan. 2011. Quality–speed conundrum: Trade-offs in customer-intensive services. *Management Science* **57**(1) 40–56.
- Arnott, R., E. Inci. 2006. An integrated model of downtown parking and traffic congestion. *Journal of Urban Economics* **60**(3) 418 – 442.
- Ashlagi, I., P. Shi. 2016. Optimal allocation without money: An engineering approach. *Management Science* **62**(4) 1078–1097.
- Bonsall, P., J. Shires, J. Maule, B. Matthews, J. Beale. 2007. Responses to complex pricing signals: Theory, evidence and implications for road pricing. *Transportation Research* **41**(7) 672 – 683.
- Chen, Y., S. Yao. 2016. "sequential search with refinement: Model and application with click-stream data". *Management Science* Forthcoming.
- De Los Santos, B., A. Hortaçsu, M.R. Wildenbeest. 2012. Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review* **102**(6) 2955–80.
- Emadi, S.M., B. Staats. 2017. A structural estimation approach to agent attrition. Working Paper, University of North Carolina, Chapel Hill, NC.
- Feldman, P., E. Segev. 2018. The important role of time limits when consumers choose their time in service. Working Paper, Boston University.

- 
- Fisher, M., S. Gallino, J. Li. 2015. Competition-based dynamic pricing in online retailing: A methodology validated with field experiments. *Management Science* Forthcoming.
- Gallien, J., I. Rashkova, R. Atun, P. Yadav. 2016. National drug stockout risks and the global fund disbursement process for procurement. *Production and Operations Management* .
- Hassin, R., M. Haviv. 2003. *To Queue or Not to Queue - Equilibrium Behavior in Queueing*. Springer, New York.
- Hong, H., M. Shum. 2006. Using price distributions to estimate search costs. *The RAND Journal of Economics* **37**(2) 257–275.
- Honka, E. 2014. Quantifying search and switching costs in the US auto insurance industry. *The RAND Journal of Economics* **45**(4) 847–884.
- Hortacsu, A., C. Syverson. 2004. Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly Journal of Economics* **119**(2) 403–456.
- Jaeker, J.A. Berry, A.L. Tucker. 2017. Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Science* **63**(4) 1042–1062.
- Kabra, A., E. Belavina, K. Girotra. 2015. Bike share systems: Accessibility and availability. Working Paper, INSEAD, Fontainebleau, France.
- Kc, D.S., C. Terwiesch. 2009. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science* **55**(9) 1486–1498.
- Kim, J.B., P. Albuquerque, B.J. Bronnenberg. 2010. Online demand under limited consumer search. *Marketing Science* **29**(6) 1001–1023.
- Kim, S-H, C.W. Chan, M. Olivares, G. Escobar. 2015. ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* **61**(1) 19–38.
- Kok, A.G., K. Shang, S. Yucel. 2016. Impact of electricity pricing policies on renewable energy investments and carbon emissions. *Management Science* Forthcoming.
- Koulayev, S. 2014. Search for differentiated products: identification and estimation. *The RAND Journal of Economics* **45**(3) 553–575.
- Lazarev, J. 2013. The welfare effects of intertemporal price discrimination: An empirical analysis of airline pricing in us monopoly markets. Working Paper, New York University, New York City, NY.
- Lederman, R., M. Olivares, G. van Ryzin. 2014. Identifying competitors in markets with fixed product offerings. Working Paper, Columbia Business School, New York, NY.
- Leslie, P. 2004. Price discrimination in Broadway theater. *The RAND Journal of Economics* **35**(3) 520–541.
- Li, J., N. Granados, S. Netessine. 2014. Are consumers strategic? structural estimation from the air-travel industry. *Management Science* **60**(9) 2114–2137.

- 
- Mak, H.-Y., Y. Rong, Z.-J. M. Shen. 2013. Infrastructure planning for electric vehicles with battery swapping. *Management Science* **59**(7) 1557–1575.
- McFadden, Daniel. 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* **57**(5) 995–1026.
- Meijer, E., J. Rouwendal. 2006. Measuring welfare effects in models with random coefficients. *Journal of Applied Econometrics* **21**(2) 227–244.
- Moon, K., K. Bimpikis, H. Mendelson. 2017. Randomized markdowns and online monitoring. *Management Science* Forthcoming.
- Ottosson, D.B., C. Chen, T. Wang, H. Lin. 2013. The sensitivity of on-street parking demand in response to price changes: A case study in Seattle, WA. *Transport Policy* **25** 222 – 232.
- Pierce, G., D. Shoup. 2013. Getting the prices right: An evaluation of pricing parking by demand in San Francisco. *Journal of the American Planning Association* **79**(1) 67–81.
- Powell, A., S. Savin, N. Savva. 2012. Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing & Service Operations Management* **14**(4) 512–528.
- Qi, W., L. Li, S. Liu, Z.-J. M. Shen. 2018. Shared mobility for last-mile delivery: Design, operational prescriptions, and environmental impact. *Manufacturing & Service Operations Management* .
- Schneider, F., U. W. Thonemann, D. Klabjan. 2017. Optimization of battery charging and purchasing at electric vehicle battery swap stations. *Transportation Science* .
- Shapiro, S.L., J. Drayer. 2014. An examination of dynamic ticket pricing and secondary market price determinants in major league baseball. *Sport Management Review* **17** 145–159.
- Shoup, D. 2005. *The High Cost of Free Parking*. Planners Press, Chicago, IL.
- Shoup, D. 2006. Cruising for parking. *Transport Policy* **13** 479–486.
- Stavins, R.N. 2011. The problem of the commons: Still unsettled after 100 years. *American Economic Review* **101**(1) 81–108.
- Tereyagolu, N., P.S. Fader, S. Veeraraghavan. 2016. Pricing theater seats: The value of price commitment and monotone discounting. *Production and Operations Management* Forthcoming.
- Tong, C., S. Rajagopalan. 2014. Pricing and operational performance in discretionary services. *Production and Operations Management* **23**(4) 689–703.
- Vickrey, W.S. 1952. The revision of the rapid transit fare structure of the city of New York. Mimeo. Technical Monograph 3. Finance Project, Mayor’s Committee for Management Survey of the City of New York.
- Vulcano, G., G. van Ryzin, W. Chaar. 2010. Choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing and Service Operations Management* **12**(3) 371–392.
- Wang, J. 2017. Extended warranty pricing with product failures and forward looking consumers: An empirical approach using sales data. Working Paper, The University of Hong Kong.

- Williamson, O.E. 1966. Peak-load pricing and optimal capacity under indivisibility constraints. *The American Economic Review* **56**(4) 810–827.
- Xu, J., P. Fader, S. Veeraraghavan. 2016. Evaluating the effectiveness of dynamic pricing strategies on mlb single-game ticket revenue. Working Paper.
- Yu, Q., G. Allon, A. Bassamboo. 2016. How do delay announcements shape customer behavior? an empirical study. *Management Science* Published online in March 2016.

## Appendix A: Calculation of Utilities and The Value of Continued Search

In this section, we derive the maximum utilities conditional on availability and show how the value of continued search can be calculated recursively.

- When the block is available, i.e.,  $A_{rtdb_k} = 1$ ,

$$\begin{aligned}
& u_i^*(b_k, P_{rtdb_k}, 1, \epsilon_{irtdb_k}) \\
&= \begin{cases} -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k), & \text{if } -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k) > \max(V_i^{\text{garage}|\text{o}}, V_i^{\text{park}}(b_k, P_{rtdb_k})), \\ V_i^{\text{park}}(b_k, P_{rtdb_k}), & \text{if } V_i^{\text{park}}(b_k, P_{rtdb_k}) \geq \max(V_i^{\text{garage}|\text{o}}, -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k)), \\ V_i^{\text{garage}|\text{o}}, & \text{otherwise.} \end{cases} \\
&= \begin{cases} -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k), & \text{if } \epsilon_{irtdb_k} < \frac{V_i^{\text{search}}(b_k) - V_i^{\text{garage}|\text{o}}}{s_i}, P_{rtdb_k} \geq \frac{s_i \epsilon_{irtdb_k} - V_i^{\text{search}}(b_k) + v_{rtd} + \epsilon_{irtd} - \eta_i d(b_i^*, b_k)}{\theta_i h_i}, \\ V_i^{\text{park}}(b_k, P_{rtdb_k}), & \text{if } P_{rtdb_k} < \frac{v_{rtd} + \epsilon_{irtd} - \eta_i d(b_i^*, b_k) - V_i^{\text{garage}|\text{o}}}{\theta_i h_i}, \epsilon_{irtdb_k} > \frac{V_i^{\text{search}}(b_k) - V_i^{\text{park}}(b_k, P_{rtdb_k})}{s_i}, \\ V_i^{\text{garage}|\text{o}}, & \text{otherwise.} \end{cases}
\end{aligned}$$

- If the block is unavailable, i.e.,  $A_{rtdb_k} = 0$ ,

$$\begin{aligned}
& u_i^*(b_k, P_{rtdb_k}, 0, \epsilon_{irtdb_k}) \\
&= \begin{cases} -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k), & \text{if } -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k) > V_i^{\text{garage}|\text{o}}, \\ V_i^{\text{garage}|\text{o}}, & \text{otherwise,} \end{cases} \\
&= \begin{cases} -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k), & \text{if } \epsilon_{irtdb_k} < \frac{V_i^{\text{search}}(b_k) - V_i^{\text{garage}|\text{o}}}{s_i}, \\ V_i^{\text{garage}|\text{o}}, & \text{otherwise.} \end{cases}
\end{aligned}$$

Next, we show that  $V_i^{\text{search}}(b_k)$  can be calculated recursively. Under the assumption that consumer beliefs of  $(P_{rtdb_k}, A_{rtdb_k})$  are i.i.d. across blocks, we can write the expected value of a consumer from continuing to search:

$$\begin{aligned}
V_i^{\text{search}}(b_k) &= E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \right] \\
&= E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | b_k \right] \\
&= E \left[ E_{(P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}})} \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_{k+1}, b_k) \right] \right] \\
&= \frac{1}{|B_{rb_k}|} \sum_{b_{k+1} \in B_{rb_k}} \left[ \iiint \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) \right. \right. \\
&\quad \left. \left. \cdot f_{rtd}^{P,A}(P_{rtdb_{k+1}}, A_{rtdb_{k+1}}) f_{rtd}^\epsilon(\epsilon_{irtdb_{k+1}}) dP_{rtdb_{k+1}} dA_{rtdb_{k+1}} d\epsilon_{irtdb_{k+1}} \right] \right] \\
&= \frac{1}{|B_{rb_k}|} \sum_{b_{k+1} \in B_{rb_k}} \left[ \phi_{rtd} \iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P, 1, \epsilon; a) f_{rtd}^{P|A}(P|A=1) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \right. \\
&\quad \left. + (1 - \phi_{rtd}) \iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P, 0, \epsilon; a) f_{rtd}^{P|A}(P|A=0) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \right] \quad (5)
\end{aligned}$$

The first double integral in the brackets represents the expected value-to-go if the next block is available, i.e.,  $A_{rtdb_{k+1}} = 1$ . The second double integral in the brackets represents the expected value-to-go if the next block is unavailable, i.e.,  $A_{rtdb_{k+1}} = 0$ . The function  $f_{rtd}^{P|A}$  denotes the density function of price conditional on availability. We expand these two double integrals further:

$$\iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P, 1, \epsilon; a) f_{rtd}^{P|A}(P|A=1) f_{rtd}^\epsilon(\epsilon) dP d\epsilon$$

$$\begin{aligned}
&= \int_{\epsilon \leq \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i}} \left( -s_i \epsilon + V_i^{\text{search}}(b_{k+1}) \right) \\
&\cdot \bar{F}_{rtd}^{P|A} \left( \frac{s_i \epsilon - V_i^{\text{search}}(b_{k+1}) + v_{rtd} + \epsilon_{irt}d - \eta_i d(b_i^*, b_{k+1})}{\theta_i h_i} \Big| A = 1 \right) f_{rtd}^\epsilon(\epsilon) d\epsilon \\
&+ \int_{P \leq \frac{v_{rtd} + \epsilon_{irt}d - \eta_i d(b_i^*, b_{k+1}) - V_i^{\text{garage}|o}}{\theta_i h_i}} V_i^{\text{park}}(b_{k+1}, P) \bar{F}_{rtd}^\epsilon \left( \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{park}}(b_{k+1}, P)}{s_i} \right) f_{rtd}^{P|A}(P|A=1) dP \\
&+ V_i^{\text{garage}|o} \bar{F}_{rtd}^{P|A} \left( \frac{v_{rtd} + \epsilon_{irt}d - \eta_i d(b_i^*, b_{k+1}) - V_i^{\text{garage}|o}}{\theta_i h_i} \Big| A = 1 \right) \bar{F}_{rtd}^\epsilon \left( \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i} \right), \quad (6)
\end{aligned}$$

where  $F_{rtd}^\epsilon(\epsilon)$  is the CDF of  $\epsilon$ ,  $F_{rtd}^{P|A}(P|A)$  is the conditional CDF of  $P$ ,  $\bar{F}_{rtd}^\epsilon(\epsilon) = 1 - F_{rtd}^\epsilon(\epsilon)$  and  $\bar{F}_{rtd}^{P|A}(P|A) = 1 - F_{rtd}^{P|A}(P|A)$ . And

$$\begin{aligned}
&\iint \max_{a=\{0,1,2\}} u_i(b, P, 0, \epsilon; a) f_{rtd}^{P|A}(P_{rtd} b_{k+1} | A=0) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \\
&= \int_{\epsilon < \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i}} \left( -s_i \epsilon + V_i^{\text{search}}(b_{k+1}) \right) f_{rtd}^\epsilon(\epsilon) d\epsilon \\
&+ V_i^{\text{garage}|o} \bar{F}_{rtd}^\epsilon \left( \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i} \right) \quad (7)
\end{aligned}$$

In sum, we see that the value  $V_i^{\text{search}}(b_k)$  is a function of the expected values to continue searching from the adjacent blocks,  $V_i^{\text{search}}(b_{k+1})$ , the expected value of parking at the adjacent blocks,  $V_i^{\text{park}}(b_{k+1}, P)$ , and the values to stop searching  $V_i^{\text{garage}|o}$ .

## Appendix B: Simulated Methods of Moments Estimation Procedure

In this section, we provide the detailed steps of implementing Simulated Methods of Moments Estimation.

**Step 1: Solve the Model.** Given a set of parameters  $\Theta$  and simulated shocks, we solve for the decision of each consumer: whether she drives, and if so, where she parks.

1. For each consumer  $i$  in region  $r$  at time  $t$  on day  $d$ , draw the ideal location  $b_i^*$  from the ideal location distribution  $\omega_{ry}(b)$ ,  $b \in B_r$ . Draw a search cost, distance disutility and price sensitivity  $(s_i, \eta_i, \theta_i)$  from the multivariate lognormal distribution  $\ln N(\mu_{s,\eta,\theta}, W_{s,\eta,\theta})$ . Finally, we need a draw of parking duration  $h_i$ . Observe from Figure 10 that the distribution of parking duration does not fit common distributions such as normal, lognormal or extreme value. Rather, there are often spikes at 30-minute, 1-hour, 2-hour, 3-hour, and 4-hour marks. Therefore, estimating the distribution of parking duration assuming it follows certain distribution is problematic. Instead, we draw parking durations from the empirical distribution and conduct sensitivity analyses around it. The details are discussed in Appendix C.4.
2. Simulate the random utility shocks  $\epsilon_{i0}, \epsilon_{ig}, \epsilon_{is}$  from i.i.d normal distributions with mean 0 and standard deviation  $\sigma$ . Also simulate a sequence of search cost shocks  $\epsilon_{irt}db_k$ ,  $k = 1, 2, 3, \dots$  from the standard log normal distribution.
3. Calculate the utility of the outside option,  $u_{i0}$ , the utility of parking at the garage,  $u_{ig}$ , and the utility of driving and starting searching on street,  $u_{is}$ . While it is straightforward to calculate the first two utilities, to calculate  $u_{is}$ , we first need to solve the dynamic search model. We explain the details below.
4. At each step of the dynamic search, the consumer can choose whether to park at the current block, continue to search or abandon searching. Given the ideal location, the parameters, and the simulated shocks, it is straightforward to calculate the utility from parking at the current block and from giving up search. However, to calculate the expected utility from continuing to search,  $V_i^{\text{search}}(b_k)$ , we need to solve Equation (5) recursively. Observe from Equations (5),

(6) and (7), that  $V_i^{\text{search}}(b_k)$  is a function of the expected utility from continuing to search at each of the nearby blocks,  $V_i^{\text{search}}(b_{k+1})$ , where  $b_{k+1} \in B_{b_r k}$ . That is, it is possible to solve for the fixed point of the vector  $V_i^{\text{search}}(b)$  using the system of  $|B_r|$  equations. Note that  $\phi_{rtd}$  and the conditional distribution  $f_{rtd}^b|A$  are calculated based on observations in the data under the rational expectation assumption. In the counterfactual analyses, however, we calculate new equilibrium distributions following the procedure outlined in Appendix D. Once we calculate the utility from continuing to search, we solve for the optimal decision at each step according to the optimal decision rule. Finally, we calculate the expected utility of a consumer who chooses to drive and start searching on street,  $u_{is}$ .

5. Determine whether a consumer decides to drive and, if so, whether she starts searching on street or parks at the garage directly.
6. Repeat these steps for all  $M_r$  consumers in the region at time  $t$  on day  $d$ .
7. Repeat these steps for all times and days in all regions.

**Step 2: Calculate Simulated Moments.** Based on the simulation, for each region, time and day, calculate 1) the number of consumers who choose to drive and end up parking at each block, 2) the number of consumers who choose to drive and end up parking at the garage, 3) the total minutes parked at each block, and 4) the total minutes parked at the garage.

**Step 3: Match Simulated Moments to Observed Moments.** Solve the moment conditions in Equation (4) to obtain the parameter estimates  $\hat{\Theta}$ .

## Appendix C: Robustness Tests

We evaluate the robustness of our model along the following dimensions:

### C.1. Market Size

In the baseline model, we set the market size to be twice the average number of drivers in the before period, which yields a market size of 700 in Marina, 1,135 in Fillmore and 1,420 in Mission. We perform the sensitivity analysis by choosing different levels of market size, which are 1.5 and 2.5 times the average number of drivers in the before period. This choice rule yields market sizes of 525 and 875 for Marina, 850 and 1,420 for Fillmore, and 1,065 and 1,775 for Mission. Results in Figure 6 shows that changes in consumer surplus, social welfare and search traffic are robust to the choice of market size.

### C.2. Consumer Beliefs

Although consumers may not have complete knowledge of the states of every individual block, it may be restrictive to assume that consumers' price and availability beliefs are identical for all blocks. We analyze an alternative partial-knowledge model. In this model, we group blocks into three levels (high, medium, and low) according to their popularity before the implementation of congestion pricing, and allow consumers to form different beliefs of price and availability for high-, medium- and low-popularity blocks,  $\{P_{rtd}^H, A_{rtd}^H\}$ ,  $\{P_{rtd}^M, A_{rtd}^M\}$ ,  $\{P_{rtd}^L, A_{rtd}^L\}$ . We also assume that consumers know which block is high-, medium-, and low-popularity block. Empirically, we determine whether a block has a high, medium, or low popularity based on SFpark's definition. High popularity blocks are those with an average occupancy rate of 80% or higher before implementation. Medium popularity blocks are those with average occupancy rates of between 60% and 80%. Low popularity blocks are those with average occupancy rates of 60% or lower.

Under such assumptions, consumers' decisions differ from those derived under identical beliefs in two ways. First, consumers may choose to start their search from a block that is not their destination given their knowledge of all blocks' popularity. Second, consumers will derive an optimal search path instead of following a random walk strategy at the intersection: they choose where to go next based on which block yields the highest expected utility. We now derive the transition probability and the optimal decision rule under such assumptions.

Let  $b_{k+1}^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k})$  denote the block that consumer  $i$  will visit if she chooses to continue searching after block  $b_k$ . Let  $U_b = \{H, M, L\}$  denote the popularity level of block  $b$ . We can now write down the transition probability:

$$\begin{aligned} & Pr(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; 0 | b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \\ &= \begin{cases} f_{rtdb}^{P,A|U_{b_{k+1}}} (P_{rtdb_{k+1}}, A_{rtdb_{k+1}} | b_k, P_{rtdb_k}, A_{rtdb_k}) f^\epsilon(\epsilon_{irtdb_{k+1}}), & \text{if } b_{k+1} = b_{k+1}^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}), \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

where  $f_{rtdb}^{P,A|U}$  is the joint density function of  $P, A$  for block type  $U$ . Under the assumptions that consumers form rational expectations for each block type and that consumers do not update their beliefs, the value of continued search can be simplified as:

$$\begin{aligned} & V_i^{\text{search}}(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \\ &= E \left[ \max_{a=\{0,1,2\}, b_{k+1} \in B_{rb_k}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \right] \\ &= \max_{b_{k+1} \in B_{rb_k}} E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_{k+1}, b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \right] \\ &= \max_{b_{k+1} \in B_{rb_k}} E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_{k+1}, b_k) \right] \\ &\equiv \max_{b_{k+1} \in B_{rb_k}} V_i^{\text{search}}(b_{k+1}, b_k) \equiv V_i^{\text{search}}(b_k). \end{aligned} \quad (8)$$

*Optimal Decision Rule.* The optimal decision rule now concerns two decisions: (1) whether to park, continue or stop searching, and (2) which block to search next if the search continues. For consumer  $i$ , the optimal decision rule  $\{a_i^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}), b_{k+1}^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k})\}$  can be characterized as follows. If consumer  $i$  decides to continue searching, she will drive to block  $b_{k+1}^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k})$  such that:

$$\begin{aligned} & b_{k+1}^*(b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \\ &= \arg \max_{b_{k+1} \in B_{rb_k}} E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_{k+1}, b_k, P_{rtdb_k}, A_{rtdb_k}, \epsilon_{irtdb_k}) \right] \\ &= \arg \max_{b_{k+1} \in B_{rb_k}} E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_{k+1}, b_k) \right] \\ &= \arg \max_{b_{k+1} \in B_{rb_k}} E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | b_k \right] \end{aligned}$$

The optimal decision  $a_i^*(b_k, P_{rtdb_k}, A_{rtdb_k})$  can be characterized similarly as before. When the current block is available, i.e.,  $A_{rtdb_k} = 1$ , a consumer will choose the action that gives her the highest utility:

$$a_i^*(b_k, P_{rtdb_k}, 1, \epsilon_{irtdb_k}) = \begin{cases} 0, & \text{if } -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k) > \max(V_i^{\text{garage}^{\text{lo}}}, V_i^{\text{park}}(b_k, P_{rtdb_k})) \\ 1, & \text{if } V_i^{\text{park}}(b_k, P_{rtdb_k}) \geq \max(V_i^{\text{garage}^{\text{lo}}}, -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k)) \\ 2, & \text{otherwise.} \end{cases}$$

When the current block is unavailable, i.e.,  $A_{rtdb_k} = 0$ ,

$$a_i^*(b_k, P_{rtdb_k}, 0, \epsilon_{irtdb_k}) = \begin{cases} 0, & \text{if } -s_i \epsilon_{irtdb_k} + V_i^{\text{search}}(b_k) > V_i^{\text{garage}^{\text{lo}}} \\ 2, & \text{otherwise.} \end{cases}$$

Note that  $V_i^{\text{search}}(b_k)$  is defined by Equation (8), and can be calculated recursively.

$$V_i^{\text{search}}(b_k) = \max_{b_{k+1} \in B_{rb_k}} E \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtdb_{k+1}}, A_{rtdb_{k+1}}, \epsilon_{irtdb_{k+1}}; a) | (b_{k+1}, b_k) \right]$$

$$\begin{aligned}
&= \max_{b_{k+1} \in B_{rbk}} \left[ \iiint \left[ \max_{a=\{0,1,2\}} u_i(b_{k+1}, P_{rtd}^{U_{b_{k+1}}}, A_{rtd}^{U_{b_{k+1}}}, \epsilon_{irtdb_{k+1}}; a) \right. \right. \\
&\quad \left. \left. \cdot f_{rtd}^{P,A|U_{b_{k+1}}} (P_{rtd}^{U_{b_{k+1}}}, A_{rtd}^{U_{b_{k+1}}}) f_{rtd}^\epsilon(\epsilon_{irtdb_{k+1}}) dP_{rtd}^{U_{b_{k+1}}} dA_{rtd}^{U_{b_{k+1}}} d\epsilon_{irtdb_{k+1}} \right] \right] \\
&= \max_{b_{k+1} \in B_{rbk}} \left[ \phi_{rtd}^{U_{b_{k+1}}} \iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P, 1, \epsilon; a) f_{rtd}^{P|A, U_{b_{k+1}}} (P|A^{U_{b_{k+1}}=1}) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \right. \\
&\quad \left. + (1 - \phi_{rtd}^{U_{b_{k+1}}}) \iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P, 0, \epsilon; a) f_{rtd}^{P|A, U_{b_{k+1}}} (P|A^{U_{b_{k+1}}=0}) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \right], \quad (9)
\end{aligned}$$

where  $\phi_{rtd}^U$  denotes consumers' belief of availability for blocks of type  $U, U = \{H, M, L\}$ . Under the rational expectation assumption, we have  $\phi_{rtd}^U = \frac{\sum_{b \in B_r^U} \phi_{rtd}^b}{|B_r^U|}$ , where  $B_r^U, U = \{H, M, L\}$  denote the set of high-, medium, and low-popularity blocks in region  $r$ .

We can further expand the first double integral in Equation (9) as follows:

$$\begin{aligned}
&\iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P^{U_{b_{k+1}}}, 1, \epsilon; a) f_{rtd}^{P|A, U_{b_{k+1}}} (P|A^{U_{b_{k+1}}=1}) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \\
&= \int_{\epsilon \leq \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i}} \left( -s_i \epsilon + V_i^{\text{search}}(b_{k+1}) \right) \\
&\quad \cdot \bar{F}_{rtd}^{P|A, U_{b_{k+1}}} \left( \frac{s_i \epsilon - V_i^{\text{search}}(b_{k+1}) + v_{rtd} + \epsilon_{irt} - \eta_i d(b_i^*, b_{k+1})}{\theta_i h_i} \mid A^{U_{b_{k+1}}=1} \right) f_{rtd}^\epsilon(\epsilon) d\epsilon \\
&+ \int_{P^{U_{b_{k+1}}} \leq \frac{v_{rtd} + \epsilon_{irt} - \eta_i d(b_i^*, b_{k+1}) - V_i^{\text{garage}|o}}{\theta_i h_i}} V_i^{\text{park}}(b_{k+1}, P^{U_{b_{k+1}}}) \\
&\quad \cdot \bar{F}_{rtd}^\epsilon \left( \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{park}}(b_{k+1}, P^{U_{b_{k+1}}})}{s_i} \right) f_{rtd}^{P|A, U_{b_{k+1}}} (P|A^{U_{b_{k+1}}=1}) dP \\
&+ V_i^{\text{garage}|o} \bar{F}_{rtd}^{P|A, U_{b_{k+1}}} \left( \frac{v_{rtd} + \epsilon_{irt} - \eta_i d(b_i^*, b_{k+1}) - V_i^{\text{garage}|o}}{\theta_i h_i} \mid A^{U_{b_{k+1}}=1} \right) \\
&\quad \cdot \bar{F}_{rtd}^\epsilon \left( \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i} \right), \quad (10)
\end{aligned}$$

where  $F_{rtd}^\epsilon(\epsilon)$  is the CDF of  $\epsilon$ ,  $F_{rtd}^{P|A, U}(P|A)$  is the conditional CDF of  $P$  for block type  $U$ ,  $\bar{F}_{rtd}^\epsilon(\epsilon) = 1 - F_{rtd}^\epsilon(\epsilon)$  and  $\bar{F}_{rtd}^{P|A, U_{b_{k+1}}}(P|A^{U_{b_{k+1}}}) = 1 - F_{rtd}^{P|A, U_{b_{k+1}}}(P|A^{U_{b_{k+1}}})$ .

We can expand the second double integral in Equation (9) as follows:

$$\begin{aligned}
&\iint \max_{a=\{0,1,2\}} u_i(b_{k+1}, P, 0, \epsilon; a) f_{rtd}^{P|A, U_{b_{k+1}}} (P|A^{U_{b_{k+1}}=0}) f_{rtd}^\epsilon(\epsilon) dP d\epsilon \\
&= \int_{\epsilon < \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i}} \left( -s_i \epsilon + V_i^{\text{search}}(b_{k+1}) \right) f_{rtd}^\epsilon(\epsilon) d\epsilon \\
&+ V_i^{\text{garage}|o} \bar{F}_{rtd}^\epsilon \left( \frac{V_i^{\text{search}}(b_{k+1}) - V_i^{\text{garage}|o}}{s_i} \right) \quad (11)
\end{aligned}$$

We can see that the value  $V_i^{\text{search}}(b_k)$  is a function of the expected values to continue searching from the adjacent blocks, which allows us to calculate  $V_i^{\text{search}}(b_k)$  by solving the equation recursively.

We re-estimate the model parameters and re-calculate welfare changes under the above consumer belief specification. Results presented in Figure 7 shows that changes in consumer and social welfare as well as search traffic are robust to consumers' having heterogeneous beliefs across blocks.

**Table 9** Updating Belief: Changes in Parameter Estimates

	Marina	Fillmore	Mission
Search cost (log-scaled) mean	0.00	0.00	0.00
Search cost (log-scaled) sd	0.00	0.07	0.00
Distance disutility (log-scaled) mean	0.00	0.00	0.00
Distance disutility (log-scaled) sd	0.00	0.00	0.00
Price sensitivity (log-scaled) mean	0.02	0.00	0.00
Price sensitivity (log-scaled) sd	0.03	0.00	-0.05
Search cost $\times$ distance disutility corr	0.00	0.10	0.09
Search cost $\times$ price sensitivity corr	0.00	0.00	0.00
Distance disutility $\times$ price sensitivity corr	-0.01	0.10	0.00
Trip valuation— $\alpha$	0.09	0.00	0.00
Trip valuation—intercept	0.10	0.00	0.00
Trip valuation—June	0.00	0.00	0.00
Trip valuation—July	0.00	0.00	0.00
Trip valuation—weekend	0.00	0.00	0.00
Trip valuation—noon	0.00	0.00	0.00
Trip valuation—afternoon	0.00	0.00	0.00
Destination Model $-\kappa_A$	0.00	0.00	0.00
Destination Model $-\kappa_B$	0.00	0.00	0.00
Destination Model $-\kappa_C$	0.00	0.00	0.00
Destination Model $-\kappa_D$	0.00	0.00	0.00

### C.3. Updating of Beliefs

Even though it is unlikely that consumers are armed with the sophistication to update their beliefs continuously, consumers may be able to update them periodically. In this subsection, we estimate a variant of the model where we allow consumers to update their beliefs once arriving at their ideal location and observing whether it is available or not. Under rational expectation, the updated belief is the expected availability of the all other blocks conditional on the ideal location’s availability, which can be calculated empirically using the panel availability data. In Table 9, we show that the parameter estimates are robust when allowing for consumer to update their beliefs.

### C.4. Parking Duration

To make sure that our approach of drawing the parking durations from the censored empirical distribution is not restrictive, we draw different sets of parking time distributions, which contain different proportions of short and long parking durations such as those shown in Figure 10. Specifically, we draw from the following sets of parking time distributions, which preserve the general shape of the empirical density function of parking durations:

1. The parking time distribution conditional on congestion being low, with 10% more weight for parking durations that are less than 1-hour.
2. The parking time distribution conditional on congestion being low.
3. The parking time distribution conditional on congestion being medium.
4. The parking time distribution conditional on congestion being high.
5. The parking time distribution conditional on congestion being high, with 10% less weight for parking durations that are less than 1-hour.

We re-estimate the model and re-calculate the counterfactual equilibrium. The results are displayed in Figure 8, and the changes in consumer surplus, social welfare and search traffic are very robust to the different parking duration distributions used. For example, in Mission, assuming consumers know which block is high-, medium-, and low-popularity block increases the welfare in the after period from 499.66 to 513.76, and increases the welfare in the before period from 458.60 to 458.79. That is, the welfare change is of 0.04%–2.82%.

## Appendix D: Algorithm for Computing the Counterfactual Equilibrium

In the counterfactual analysis, we compute the counterfactual equilibrium following the procedure below:

1. For each region  $r$ , time  $t$  and day  $d$ , start with an initial guess of each consumer’s driving and parking decision. Compute the number of consumers parking at each block,  $q_{rtd}(b), \forall b \in B_r$ ,

- and the total parking minutes at each block,  $Q_{rtd}(b), \forall b \in B_r$ . Also compute the number of consumers parking at the garage,  $q_{rtdg}$ , and the total minutes parked,  $Q_{rtdg}$ .
2. Calculate the occupancy rate at each block for region  $r$ , time  $t$  and day  $d$ . Then compute the availability at each block  $\hat{\phi}_{rtdb}, \forall b \in B_r$  by solving Equation (3).
  3. Given the estimated parameters and the availability vector  $\hat{\phi}_{rtdb}, \forall b \in B_r$ , re-calculate each consumer's driving and parking decision based on the optimal decision rule derived in the paper. Update  $q'_{rtd}(b), Q'_{rtd}(b), \forall b \in B_r, q'_{rtdg}$ , and  $Q'_{rtdg}$ .
  4. Repeat the above steps until convergence, i.e., 
$$\begin{cases} q'_{rtd}(b) - q_{rtd}(b), \forall b \\ Q'_{rtd}(b) - Q_{rtd}(b), \forall b \\ q'_{rtdg} - q_{rtdg}, \\ Q'_{rtdg} - Q_{rtdg} \end{cases} < \epsilon = 10^{-3}.$$
  5. Repeat the above steps to obtain the equilibrium parking locations for all regions, times and days.

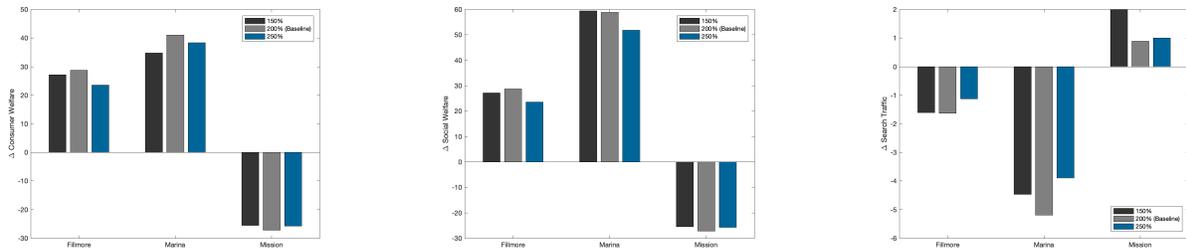


Figure 6 Robustness Test: Market Size

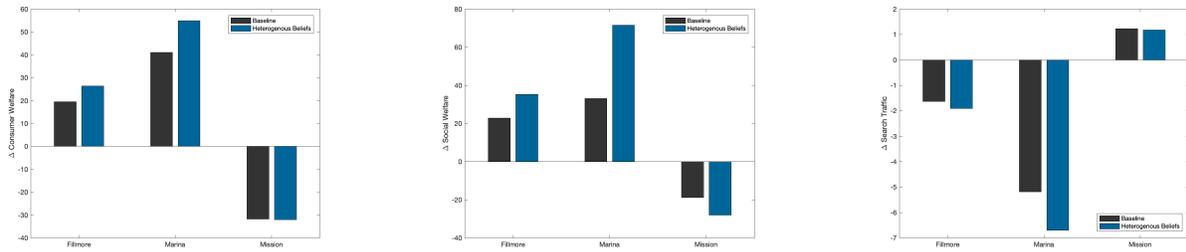


Figure 7 Robustness Test: Consumer Belief

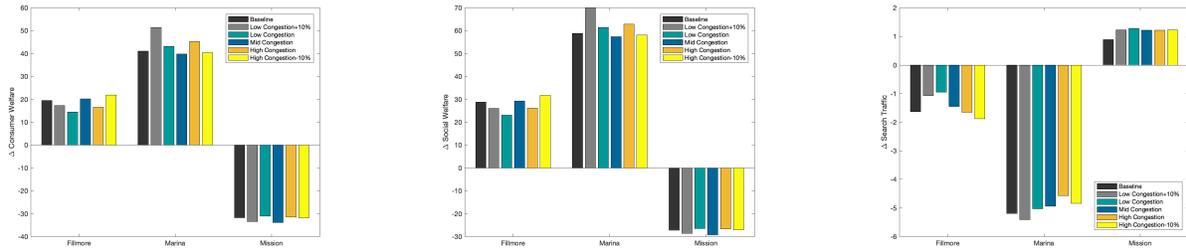


Figure 8 Robustness Test: Parking Duration

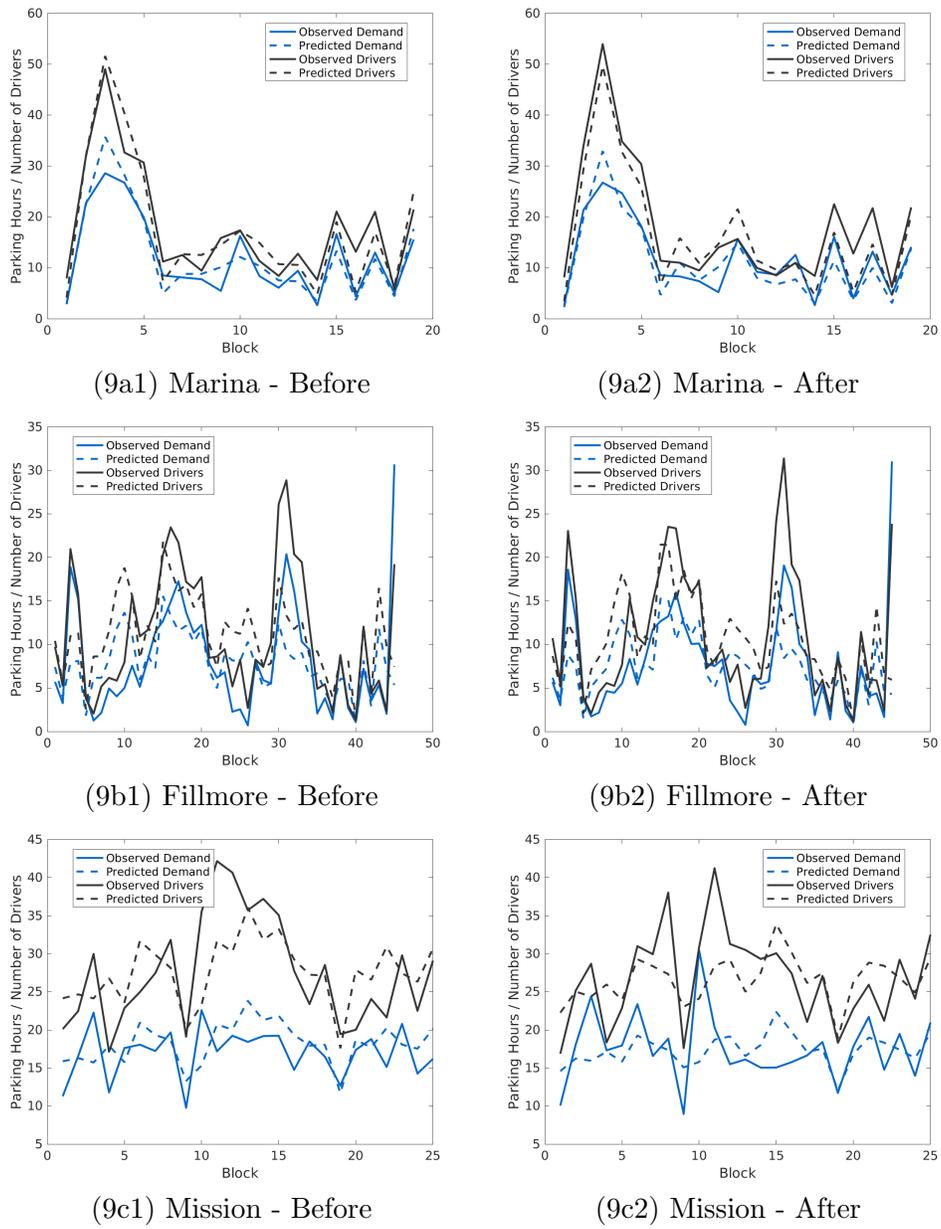


Figure 9 Moment Fits

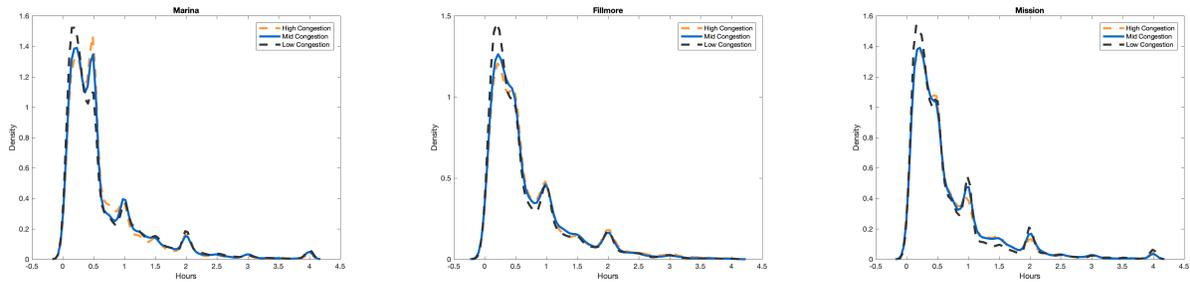


Figure 10 Parking Time Distribution