Review of Economic Studies (2023) **90**, 3186–3237 https://doi.org/10.1093/restud/rdad014

Advance access publication 18 February 2023

Left-Digit Bias at Lyft

JOHN A. LIST

Department of Economics, University of Chicago, Australian National University, ex-Lyft Chief Economist, and NBER

IAN MUIR

ex-Director of the Department of Economics, Lyft

DEVIN POPE

Booth School of Business, University of Chicago and NBER

and

GREGORY SUN

Olin Business School, Washington University in St. Louis

First version received July 2021*; Editorial decision August* 2022*; Accepted February* 2023 (Eds.)

Left-digit bias (or 99-cent pricing) has been discussed extensively in economics, psychology, and marketing. Despite this, we show that the rideshare company, Lyft, was not using a 99-cent pricing strategy prior to our study. Based on observational data from over 600 million Lyft sessions followed by a field experiment conducted with 21 million Lyft passengers, we provide evidence of large discontinuities in demand at dollar values. Approximately half of the downward slope of the demand curve occurs discontinuously as the price of a ride drops below a dollar value (*e.g.* \$14.00 to \$13.99). If our short-run estimates persist in the longer run, we calculate that Lyft could increase its profits by roughly \$160M per year by employing a left-digit bias pricing strategy. Our results showcase the robustness of an important behavioral bias for a large, modern company and its persistence in a highly competitive market.

Key words: Law of demand, Left-digit bias, Behavioral IO, Field experiment

JEL codes: D90, C93, D01

1. INTRODUCTION

In the past several decades, within economics, there has been a landmark shift from analysis of the market to the individual. In doing so, considerations from behavioral economics have taken a more central role. Indeed, in the 1950s, the rationality assumption inherent in economic modeling was brought into the cross-hairs by behavioral scientists, in particular, Simon's (1957) work on bounded rationality. In a seminal study that represented the neoclassical response to the behavioral movement, Becker (1962) used theory to show that the Law of Demand held at the market level, even under various irrationality ("behavioral") assumptions at the individual level. This research served to set the stage for a line of inquiry revolving around measuring the nature and extent of behavioral anomalies and how they affect market outcomes.

In this paper, we focus on one particular behavioral anomaly, left-digit bias, and show that it has important implications for how the Law of Demand actually works at the market level. The setting of our study is Lyft, the popular ride-sharing service. When seeking a ride, customers

The editor in charge of this paper was Nicola Gennaioli.

input an origin and destination and Lyft offers a price that can be accepted ("a conversion") or rejected. Our analysis focuses on cities in the USA where Lyft offers two primary products: a Standard product (one passenger is matched with one driver) and a lower-priced Shared/Pool product (potentially multiple passengers are matched with one driver). Lyft employs a complex algorithm to determine prices for these two products. Useful for our context, Lyft also created random variation in price for approximately 10% of sessions in our sample.

We begin by analyzing 7 months of observational data from 2019 in markets across the USA. During this period, we observe over 600 million price offers with a conversion rate of roughly 65%. Despite the growing literature on the left-digit bias (a literature that we review at the end of this introduction), we show that Lyft's machine learning pricing tools were not utilizing a 99-cent pricing strategy prior to our study. Specifically, Lyft was just as likely to offer a ride of \$14.00 or \$14.01 as it was to offer a ride of \$13.99 or \$13.98. Using the 10% of rides where Lyft infused random variation, we find that the Law of Demand holds. Specifically, randomly increasing the Standard (Shared) price of a ride by 1% leads to a reduction in Standard (Shared) conversion by 0.9%–1.4% (1.7%–3.3%). We also estimate cross-price elasticities which yield qualitatively expected results: the Standard product and the Shared product are substitutes.

And yet, while the Law of Demand holds, these elasticity estimates mask an important underlying behavioral feature predicted by a model of left-digit bias. We provide graphical evidence that demand curves do not show a steady and smooth decline, but that they drop discontinuously as prices cross dollar thresholds. As an example, during the 7-month period we analyze, Lyft offered 1.7 million passengers a ride for a price somewhere between \$10.96 and \$11.03 (with prices approximately equally distributed across each cent value). What were the conversion rates for those price offers? The conversion rates for \$10.96, \$10.97, \$10.98, and \$10.99 were 50.2%, 50.4%, 50.2%, and 50.1%, respectively. In contrast, the conversion rates for \$11.00, \$11.01, \$11.02, and \$11.03 were 48.7%, 48.6%, 48.7%, and 48.8%, respectively. This is not an isolated incident, but rather a pattern of discontinuous drops in demand that we document graphically across every dollar threshold.

In addition to discontinuities in own-price elasticities at dollar thresholds, we see discontinuities in cross-price elasticities. For example, moving from a Standard price of \$10.99 to \$11.00 lowers Standard conversion and overall conversion, but discontinuously increases Shared conversion. Similarly, as Shared prices move from just below to just above a dollar value, Standard conversion rates discontinuously increase.

Using a simple model of left-digit bias, we estimate a parameter value associated with the level of inattention to prices that manifests itself in this market. We estimate an inattention parameter for both Standard and Shared prices of approximately 0.5. This is true when we estimate the inattention parameter using the own-price elasticity effects as well as the cross-price elasticity effects. This parameter value indicates that passengers perceive a price that is lowered 1 cent below a dollar value as if the price was lowered by 50 cents. Put another way, the contribution of left-digit bias to the Law of Demand is such that roughly 50% of the decline in demand occurs discontinuously at dollar values.

We complement our observational analysis with a large-scale natural field experiment. For two months, starting in November of 2019, we ran a natural field experiment with 21 million users on the Lyft platform. Users were randomly placed into one of six experimental conditions that manipulated the Standard and/or Shared prices. The typical condition lowered prices that otherwise would have been between \$X.00 and \$X.09 to the 99-cent mark below \$X. Based on parameter estimates from the observational data, forecasts suggested the most effective treatments could increase profit per user by 0.3%–0.4%.

Consistent with our forecasts, we find that lowering prices from a bit above to just below a round number leads to a significant increase in profit and rides accepted. Interestingly, users that see 99-cent prices as opposed to prices that are just over the dollar amount are also more likely to increase their usage of the Lyft app in the future—which further increases profits. Aggregating up, we find that the simple pricing change that we enacted is worth approximately 0.39% in profit per user. However, given the large degree of inattention, a more aggressive pricing response to the left-digit bias of customers is optimal. Given the inattention parameter and elasticity estimates, we argue that to maximize profits all prices should be raised or lowered to a 99-cent mark, *ceteris paribus*. If this were to be done and assuming the effects we estimate hold in the long run, we estimate that Lyft could increase profits by $.25 per ride. Given ridership in 2019, a left-digit pricing strategy could increase profits by approximately $160M per year.

Using variation imposed by our field experiment, we once again estimate the level of inattention to prices in this domain. We continue to find an inattention parameter of approximately 0.5. Further, we find a remarkable level of stability in the inattention parameter when measuring its value across passengers and ride types. Even though the elasticities for different types of rides (*e.g.* airport rides versus non-airport rides) are very different, the level of inattention we observe is quite similar.

We view our findings relating to several branches of the literature. First, our paper speaks to recent work that uses large economic datasets to test for left-digit bias. Specifically, Strulov-Shlain (2021) analyzes scanner data for thousands of products and provides both reduced-form and structural evidence of left-digit bias in perceived prices by customers. He finds an inattention parameter of 0.15–0.25 and further shows that while firms do respond to this bias and price many items at the 99-cent mark, they do not respond as much as they should. A related paper, Lacetera *et al.* (2012), documents left-digit bias to odometer values in used cars. Analyzing data from millions of used-car auctions, the authors find that cars with odometer values just under a 10,000-mile threshold (*e.g.* 69,800 miles) sell for $150–$200 more than cars just over a 10,000-mile threshold (*e.g.* 70,100 miles). These discontinuous changes in value correspond to an inattention parameter estimate of 0.30.

In addition to the obvious similarities, our findings differ from these two important papers in several ways. First, the amount of data and level of statistical power in our observational analysis is enormous, even when compared to these other big data projects. In addition, the identification strategy for our observational analysis is especially clean. For example, both of the above-cited papers must grapple with excess volume of products/cars that have a price/odometer value just under the round numbers, which makes comparing the demand for products/cars that are just over versus just under a round number more difficult. Our analysis, on the other hand, is akin to a clean regression discontinuity design with a smooth density function through the thresholds. Our data also confer the ability to explore the substitution effects of left-digit bias from one product to another in a way that has previously been unexplored. Not only can we identify inattention using cross-price elasticities, but we can fully document the value to the firm of 99-cent pricing while taking into account spillover effects onto other products. Lastly, and perhaps most importantly, our paper goes beyond observational analysis and conducts a large-scale natural field experiment with millions of observations. The field experiment allows us to validate the results found in the observational data and also speak to additional key aspects of left-digit bias (*e.g.* what are the effects of 99-cent pricing on long-run behavior and platform use of customers).

Beyond the two papers singled out above, our paper adds to an even larger literature (often involving the fields of psychology and marketing) that has explored left-digit bias. For example, experimental papers have tested underlying mechanisms for the bias (see Carver and Padgett, 2012 for a discussion of nearly a dozen papers that address this issue). Other researchers have empirically estimated the impact of left-digit bias in economic markets using field experiments (*e.g.* Anderson and Simester, 2003; Ashton, 2014) and observational data (*e.g.* Stiving and Winer, 1997; Englmaier *et al.*, 2017; Hilger, 2018; Meng, 2020; Repetto and Solis, 2020). In addition, a few papers have explored theoretical implications of left-digit bias (Basu, 1997, 2006; Stiving, 2000). Aside from the literature directly testing for left-digit bias, there are additional papers that are tangentially related to left-digit bias. For example, the impact of 99-cent pricing on the efficiency and incidence of taxes (Conlon and Rao, 2020) and wage bunching at round numbers (Dube, Manning, and Naidu, 2020).

Our paper also contributes to the broader literature on behavioral industrial organization. This literature explores how firm strategies optimally account for behavioral consumers (*e.g.* DellaVigna and Malmendier, 2004) and also how firms can fail to fully optimize (Bloom and Van Reenen, 2007; Hortaçsu and Puller, 2008; Goldfarb and Xiao, 2011; Massey and Thaler, 2013; Hanna *et al.*, 2014; DellaVigna and Gentzkow, 2019). Our paper adds to this literature by providing an example of a high-tech and modern company with sophisticated pricing algorithms that failed to account for an important and well documented anomaly. As we explain in more detail when discussing our results, a number of subtle institutional and historical details about Lyft's pricing system likely biased them against finding an effect. The paper therefore provides a convincing argument for how behavioral economics matters in large and important economic markets.

Finally, to complement our reduced-form results we include structural estimates that provided insights into the degree of inattention in our data. In this manner, our study is part of a recent trend to combine a structural approach with a corresponding closely linked field experiment (Jin *et al.*, 2010; DellaVigna *et al.*, 2012; Dellavigna *et al.*, 2016a). In doing so, we highlight how this approach complements the literature on structural behavioral economics (Conlin *et al.*, 2007; Laibson *et al.*, 2007; Chetty *et al.*, 2009; Gerritsen, 2015; Dellavigna *et al.*, 2016b). As such, we show how leveraging insights from psychology and economics can be used to drive a key pricing decision within a firm, yielding a substantial profit increase.

The remainder of our paper proceeds as follows. Section 2 describes the Lyft platform and the data. Section 3 provides a theoretical framework for understanding left-digit bias in our setting. Section 4 presents the reduced-form and structural results from the observational analysis. Section 5 describes our field experiment and accompanying results. Section 6 discusses how our inattention estimates compare to other attempts at identifying an inattention parameter in the literature. In Section 7, we conclude.

2. DATA AND INSTITUTIONAL SETTING

The institutional setting of our study is the US-based company Lyft. Lyft provides rideshare services in hundreds of cities nationwide and competes directly with Uber. It is important to understand the specifics of how passengers are connected with drivers on this platform. A passenger considering a Lyft trip opens the app and inputs an origin and a destination to request a ride. Lyft calculates a price, in part, based on its estimates of the distance and duration of the trip. These events are called "sessions". For each session, Lyft records the price shown to the passenger,[1] the time when the session started, the origin and destination of the trip, a unique identifier for who is requesting the session, and whether a conversion occurred (the passenger accepted the price offered for the ride).

Importantly, when a passenger inputs an origin and destination, Lyft offers multiple products with different prices. The main product Lyft offers is its Standard ride, which matches a single

1. In the rare case when a session lasts sufficiently long, it is possible for a single session to contain multiple prices if, for example, estimates of duration change from one minute to the next based on updating traffic conditions. In these cases, Lyft records the last price shown to the passenger in the session.

TABLE 1
Summary statistics

	Standard only	Standard and shared
Average standard price	\$16.89	\$16.37
Average shared price		\$12.66
Standard conversion	0.64	0.46
Shared conversion	0.00	0.19
User had 10+ rides	0.32	0.37
User one product		0.51
User had business account	0.04	0.07
Originated in airport	0.04	0.04
Weekday session	0.67	0.70
6AM–9PM Session	0.78	0.80
Number of sessions	273M	339M

Notes: This table shows summary statistics for the data used in the observational section. The sample consists of all sessions from February 2019 through August 2019 inclusive. The unit of analysis is the session. In the left column, we subset to sessions that did not offer Shared rides. In the right column, we subset to sessions that offered both Standard and Shared rides, but where Shared Saver was unavailable. In rows "User Had 10+ Rides" through "6AM–9PM Session" and below, we show the proportion of sessions that satisfy the criterion implied by the row name. As in the text, "User One Product" is defined as the user converting to the same product (either Standard or Shared) 90+% of the time in the 28 days prior to the session.

passenger to a single nearby driver. In addition to this product, Lyft also offers a Shared product, which gives a discounted price, in exchange for allowing the app to match passengers with other passengers going in similar directions to save on costs. At the time, in a smaller number of markets, Lyft offered additional products (*e.g.* Shared Saver rides). However, we focus on the Standard and Shared products in our observational analysis given they comprised the lion's share of Lyft ridership during our sample and because those are the two products for which we randomize prices in our natural field experiment.

The prices that Lyft offers for each of their products is based on a complicated set of algorithms. One useful feature that we will exploit in our analysis is that Lyft randomly varied the Standard and Shared prices for roughly 10% of sessions in our sample. This allows us to estimate precisely own- and cross-price elasticities for Standard and Shared rides.[2]

Our analyses span two distinct time periods. First, we perform an observational analysis that utilizes data for all sessions that occurred between February and August of 2019. Summary statistics for the approximately 612 million sessions during this period are presented in Table 1. As can be seen in the table, about half of the sessions occurred in markets offering only the Standard product (273M) and the other half occurred in markets offering both the Standard and Shared products (339M). The average Standard price offered was between \$16 and \$17 with the average Shared price being approximately \$4 cheaper. A conversion (the passenger accepts a price offer on either Standard or Shared) occurred between 60% and 70% of the sessions. When both Standard and Shared prices were offered, the Standard product was more than twice as likely to be chosen as the Shared product. The summary statistics table also contains additional information about the passengers and requested rides.

2. For the randomized price sessions, Lyft drew two random multipliers ranging from 0.8 to 1.2. The Standard price is multiplied by the first multiplier and the Shared price is multiplied by the second multiplier. As researchers, we know which rides were randomized and the multiplier on each. One additional complexity in determining Shared price is that business logic at the time dictated that Shared prices should be at least 10% cheaper than Standard prices. When binding, this adjustment to Shared prices is made.

The second time period we analyze is November 18, 2019 to January 22, 2020. During this time period, we conduct a nationwide experiment on the Lyft platform making use of observations from more than 21 million Lyft riders whose prices were systematically varied to dig deeper into the insights gained from the observational data. The field experiment and resulting data are described in more detail below.

3. THEORETICAL FRAMEWORK

In this section, we describe the theoretical framework through which we interpret our evidence of left-digit bias. In Section 3.1, we present a basic model in a setting where we posit a demand curve without explicit "microfoundations" and assume away heterogeneity in inattention. This allows us to highlight the key intuitions of our left-digit bias framework while abstracting away from many of the technical details that are necessary to fully connect our theoretical framework to the specific institutional details of our empirical setting. In Section 3.2, we show that most of the insights from our basic model, with some slight modifications, continue to hold when we explicitly model the discrete choice context of Lyft's data and allow for heterogeneity in the inattention parameter. The reader who is uninterested in the technical details may skip much of Section 3.2.

3.1. *Basic model*

We model left-digit bias through the lens of inattention. Our basic model follows the framework laid out in Chetty *et al.* (2009) and DellaVigna (2009), where consumers pay full attention to a visible component of a good and are somewhat inattentive to an opaque component of the same good.[3] To keep the exposition simple, in this section, we will abstract from some of the finer details relevant to our empirical setting and describe our model for a representative agent. In the next section, after describing the institutional details of our Lyft setting, we will describe a more empirically relevant model which we explicitly connect to our data. The key results gleaned from the model described in this section largely carry through to our empirical model, but the reader may find the discussion here helpful for gaining intuition in a simpler context.

We begin by describing our model for the case of a single good. We model that consumers behave in accordance with a latent demand curve $D^*(p)$ and systematically misperceive prices by not fully internalizing the cent amount in the price.[4] The parameter $\theta \in [0, 1]$ represents *in*attention

$$\underbrace{\hat{p}(p;\theta)}_{\text{Perceived Prices}} = \underbrace{\lfloor p \rfloor}_{\text{Dollars}} + \underbrace{(1-\theta)}_{\text{Attention}} \times \underbrace{(p - \lfloor p \rfloor)}_{\text{Cents}} \tag{1}$$

3. The predictions that we produce in this section are similar to those produced by Lacetera *et al.* (2012) and Strulov-Shalin (2021). We focus, however, specifically on predictions of the model that translate to our particular institutional setting (*e.g.* a setting with two products).

4. Most models of left-digit bias struggle to fully capture what happens when the number of digits changes (*e.g.* a product's price moves from \$9.99 to \$10.00). One option is to have an inattention parameter for the first digit and a second inattention parameter for the second digit of a price. Alternatively, the modeler may choose in any given context which numbers are "visible" and which are "opaque". In Lacetera *et al.* (2012) and Busse *et al.* (2013), they argue that the visible component of, for example, 25,623 miles on an odometer is 20,000 miles and the visible component of, for example, 122,367 miles on an odometer is 120,000 miles. Strulov-Shlain (2021) argues that the dollar amounts of a product are visible and the cent values are opaque. We follow that same path here; specifically, we argue that the cent values are opaque and the dollar amount is the "left-digit" (even if the left-digit is a two-digit number). Thus, consumers in our theory perceive the difference between the prices \$9.99 and \$10.00 in the same way that they perceive the difference between the prices \$13.99 and \$14.00.

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. For example, under this model, if a consumer sees a price of \$18.42 and has an inattention parameter of $\theta = 0.5$, then they mistakenly perceive the price to instead be $\$18.21 = 18.00 + 0.5 \cdot (0.42)$. Accordingly, observed demand is related to inattention and latent demand by

$$D(p;\theta) = D^*(\hat{p}(p;\theta)) \tag{2}$$

We make the following assumption for convenience when stating our formal results:

Assumption 1. $D^*(p)$ is linear, that is, $D^*(p) = \alpha - \beta p$.

Proposition 1. *When consumers are inattentive ($\theta > 0$) and demand curves slope downward ($\beta > 0$), demand curves are discontinuous around integer prices.*

For completeness, we provide proofs of propositions stated in the main text in Appendix A. We can characterize this discontinuity quantitatively through the model, but to do so, we need a measure of "price sensitivity" through which to compare the effects of different price changes. Given Assumption 1, we use the following definition of price sensitivity:

Definition 1. Let p^{new}, p^{old} be two prices. Then, the *average slope* of the demand curve between p^{new} and p^{old} is given by $\frac{D(p^{new};\theta)-D(p^{old};\theta)}{p^{new}-p^{old}}$.

Consider now, the case where we compare prices around a single dollar amount, so $p^{new} = \$ + ¢_1$ and $p^{old} = \$ - ¢_2$, where \$ is an integer and $¢_i \in [0,1)$. Under our model of inattention, a price change is perceived as a change of $\$ + (1-\theta)¢_1 - (\$ - 1 + (1-\theta)\cdot(1 - ¢_2)) = (1-\theta)\cdot(¢_1 + ¢_2) + \theta$, so $D(p_1;\theta) - D(p_2;\theta) = \beta[(1-\theta)\cdot(¢_1+¢_2)+\theta]$. Alternatively, the actual difference in prices is $-(¢_1 + ¢_2)$, so the average slope of the demand curve between p_1 and p_2 is

$$-\beta\frac{(1-\theta)\cdot(¢_1+¢_2)+\theta}{¢_1+¢_2}$$

In the special case where $¢_1 + ¢_2 = 1$, (*i.e.* when the cents amount in the old and new prices are exactly equal), the perceived and actual price changes are identical; thus, in this case, the average slope of the demand curve is $-\beta$, the slope of the latent demand curve. Comparing this benchmark case to a price change where the final cents amount varies therefore provides us with an approach to identify θ. We summarize this discussion in the following proposition.

Proposition 2. *Suppose latent demand is linear as in Assumption 1 and let $¢_1, ¢_2, ¢_3 \in [0,1)$. Let \$ be an integer price. Let s_1 be the average slope between prices $\$ + ¢_1$ and $\$ - ¢_2$ and let s_2 be the average slope between $\$ + ¢_3$ and $\$ - 1 + ¢_3$. Then $s_1/s_2 = \frac{(1-\theta)\cdot(¢_1+¢_2)+\theta}{¢_1+¢_2}$. In particular, $\lim_{¢_1,¢_2\to 0} s_1/s_2(¢_1+¢_2) = \theta$.*

This proposition makes it clear that the price sensitivity measured from examining price changes just crossing dollar amounts is larger than the price sensitivity measured when comparing larger price changes. Figure 1 shows the intuition for the above result. This result, and its later extensions, provide the key motivation for our empirical strategies throughout the paper: each time we estimate an inattention parameter, we are doing so by taking a ratio between a discontinuity and a slope of a demand curve, although the statistical details in how we implement this idea vary depending on the particular dataset.

If the latent demand curve is not linear but close to linear, then dividing the size of the discontinuity by the average slope of the demand curve around a dollar change approximates the inattention parameter. The formal justification for this statement is given by the below result.

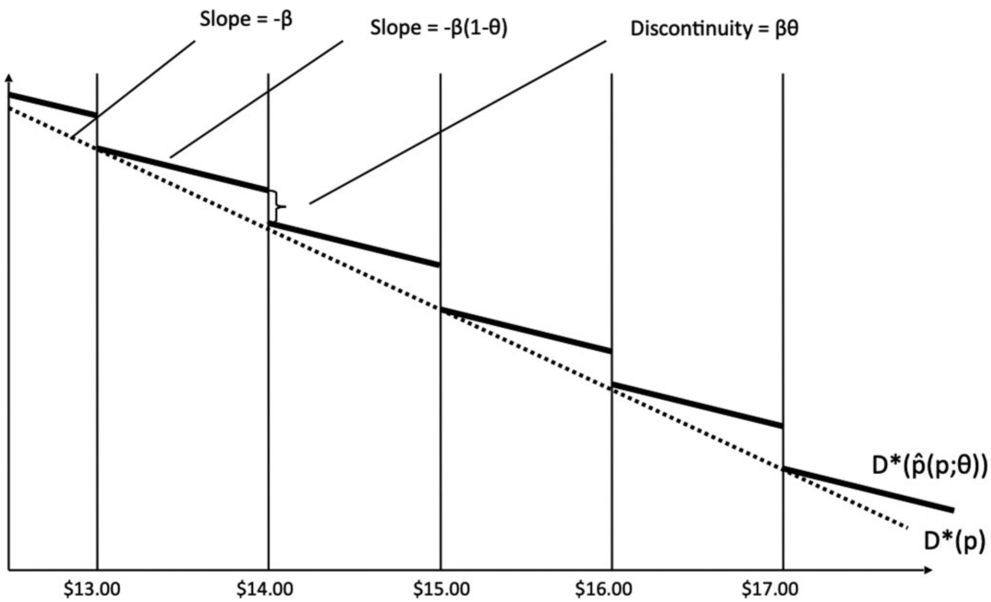


FIGURE 1
The inattention model

Notes: This figure shows our model of inattention. The dotted line represents the latent demand curve while the solid line is the demand consumers actually follow.

Proposition 3. *Assume* $\theta > 0$. *Let* $¢_1, ¢_2, ¢_3 \in [0, 1)$. *Let* $\$$ *be an integer price. Let* s_1 *be the average slope between prices* $\$ + ¢_1$ *and* $\$ - ¢_2$ *and let* s_2 *be the average slope between* $\$ + ¢_3$ *and* $\$ - ¢_3$. *Then,*

$$\lim_{¢_1, ¢_2 \to 0} s_1/s_2(¢_1 + ¢_2) = \theta \cdot \frac{\frac{\int_{\$-\theta}^{\$} (D^*)'(t)\,\mathrm{d}t}{\theta}}{\int_{\$-(1-\theta)¢_3-\theta}^{\$+(1-\theta)¢_3} (D^*)'(t)\,\mathrm{d}t}$$

The numerator and the denominator are averages of $(D^*)'$ over different ranges. Importantly, the range of integration in both cases is limited to being within a dollar of the integer price, $. This implies that provided D^* is sufficiently smooth in the sense that $(D^*)'$ does not vary too dramatically within the dollar amount, the fraction should be close to 1, so the RHS will be close to the inattention parameter θ. In what follows, we continue to frame our discussion in terms of linear latent demand curves, but the reader should keep in mind that the idea behind Proposition 3 implies that our main results will be fairly robust to mild deviations from the linearity assumption.

Our inattention model is useful in organizing the data and provides a strong, but qualitative testable, prediction: discontinuities in demand will exist around dollar values if $\theta > 0$ and $\beta > 0$. When we extend the model to multiple goods, we obtain an additional, more quantitative, testable prediction. For ease of exposition (and because it mirrors our setting at Lyft), consider the case of two goods. Let the two goods be indexed $i = 1, 2$ and posit a pair of latent demand curves $D_i^*(p_1, p_2)$. We continue to assume that people deviate from the latent demand curves because they systematically misperceive prices

$$\begin{aligned}\hat{p}_1(p_1;\theta_1) &= \lfloor p_1 \rfloor + (1-\theta_1)\cdot(p_1 - \lfloor p_1 \rfloor)\\ \hat{p}_2(p_2;\theta_2) &= \lfloor p_2 \rfloor + (1-\theta_2)\cdot(p_2 - \lfloor p_2 \rfloor)\end{aligned} \tag{3}$$

Where θ_1, θ_2 are, respectively, the inattention to the price of good 1 and good 2. Once again, demands for good 1 and good 2 are related to inattention and the latent demand curves by

$$\begin{aligned}D_1(p_1,p_2;\theta_1,\theta_2) &= D_1^*(\hat{p}_1(p_1;\theta_1),\hat{p}_2(p_2;\theta_2))\\ D_2(p_1,p_2;\theta_1,\theta_2) &= D_2^*(\hat{p}_1(p_1;\theta_1),\hat{p}_2(p_2;\theta_2))\end{aligned} \tag{4}$$

Modeling left-digit bias in this manner corresponds to a meaningful restriction on behavior consistent with the model. Specifically, implicit in the above formulation is that the effect of left-digit bias is entirely mediated through its effect on perceptions of price. To see why this restricts the set of possible data consistent with the model, we assume an analog of Assumption 1 and show how it can be used to extend Proposition 2.

Assumption 2. $D_1^*(p_1,p_2)$ and $D_2^*(p_1,p_2)$ are linear, that is $D_1^*(p_1,p_2) = \alpha_1 - \beta_{1,1}p_1 + \beta_{1,2}p_2$ while $D_2^*(p_1,p_2) = \alpha_2 + \beta_{2,1}p_1 - \beta_{2,2}p_2$.[5]

Definition 2. Let p_i^{new}, p_i^{old} be two prices for good i. Then, the *average partial slope* of good j's demand curve with respect to the price of good i between p_i^{new} and p_i^{old} is given by $\frac{D_j(p_i^{new},p_{-i};\theta)-D_j(p_i^{old},p_{-i};\theta)}{p_i^{new}-p_i^{old}}$ where p_{-i} is the price of the other good, which is held fixed.

Proposition 4. *Suppose latent demand is linear as in Assumption 2 and let* $¢_1, ¢_2, ¢_3 \in [0,1)$. *Let* \$ *be an integer price. Let* $s_{ji,1}$ *be the average partial slope of good* j*'s demand curve with respect to the price of good* i *between prices* $\$+¢_1$ *and* $\$-¢_2$ *and let* $s_{ji,2}$ *be the average partial slope of good* j*'s demand curve with respect to the price of good* i *between* $\$+¢_3$ *and* $\$-1+¢_3$. *Then* $s_{ji,1}/s_{ji,2} = \frac{(1-\theta_i)\cdot(¢_1+¢_2)+\theta_i}{¢_1+¢_2}$. *In particular,* $\lim_{¢_1,¢_2\to 0} s_{ji,1}/s_{ji,2}(¢_1+¢_2) = \theta_i$.[6]

Note how in the above proposition, the slope ratio is defined separately for each good j whose demand we are measuring and each good i whose price we are varying. However, the identity of good j completely drops out of the ratio formula. This reflects an additional testable restriction of the inattention model when there are multiple goods because it implies that from a single cause (inattention to one price), there are quantitatively related effects. We formalize this intuition as follows:

Corollary 1. *Suppose latent demand is linear as in Assumption 2 and let* $¢_1, ¢_2, ¢_3 \in [0,1)$. *Let* \$ *be an integer price, and let* i, j, k *be goods such that* $j \neq k$. *Define* $s_{ji,1}, s_{ji,2}, s_{ki,1}, s_{ki}$ *as in Proposition 4. Then*

$$\frac{s_{ji,1}}{s_{ji,2}} = \frac{s_{ki,1}}{s_{ki,2}}$$

We provide a simple example to elucidate the economic content of the above restriction.

5. In the same spirit as Proposition 3, we may view the linearity assumption made here as an approximation to the true latent demand curve. The results presented below are therefore robust to mild deviations from linearity.

6. Note that this proposition would continue to hold even if we allowed a mild non-linearity to the demand curve in the form of an interaction term between p_1 and p_2. To see why, note that this result is stated in terms of holding one of the prices fixed and varying the other price. In the presence of an interaction term, the demand curve is still linear in the price which is being varied, holding the other price fixed, the conclusion of the Proposition will continue to hold.

Example 1. Consider two goods and fixed price points p_1, p_2. Suppose that when the price of good 1 rises by \$1, demand for good 2 rises by 30% of the amount by which demand for good 1 decreased. This tells us that among the population of people who were roughly indifferent (more precisely, with a reservation value for good 1 compared to the next best option within \$1 above p_1) between good 1 and the next best option at prices p_1, p_2, 30% of them had good 2 as their next best option. Our inattention model asserts that left-digit bias affects demand *only* through its effect on perceived prices. Thus, the effect of a small price change that crosses a dollar threshold will still identify the size of the sub-population that is nearly indifferent (although the exact definition of "nearly indifferent" now changes) between good 1 and the next best option while the cross-price elasticity continues to identify the sub-population of people that is nearly indifferent between good 1 and the next best option *and* for whom good 2 is the next best option. When the distribution of preferences is sufficiently smooth (as formalized in Corollary 1 by the assumption of linear demand), we should continue to find that roughly 30% of the reduction in demand in good 1 is from people substituting to good 2.

Before turning to our analysis of the data, we adapt the key results from our conceptual framework to more closely match the institutional details of the market and Lyft's pricing therein.

3.2. *Inattention in a discrete choice setting*

In this subsection, we adapt the model of inattention introduced in our conceptual framework to make it more empirically relevant to the institutional details of Lyft's marketplace. The key intuitions from our conceptual framework carry over to the model described here. Here, we summarize the basic implications of the analysis in this section for readers who wish to skip to our empirical results.

Result 1. *Dividing the size of the discontinuity at integer prices by the slope of the latent demand curve at those prices identifies the average inattention of the marginal consumer (i.e. the consumer who is just indifferent between purchasing/not purchasing).*

Result 2. *Under suitable restrictions on the heterogeneity of the inattention parameter, if the inattention model describes the data generating process, the inattention measured across different consumption margins (i.e. using own- vs. cross- price effects) will be identical.*

Throughout, we adopt the usual convention of letting capital letters denote random elements while letting lower-case letters denote realizations of random variables or other non-stochastic values. Additionally, because we will be repeatedly dealing with dollars and cents, we introduce the following notation

$$\$p = \lfloor p \rfloor, \quad ¢p = p - \lfloor p \rfloor \tag{5}$$

We index which ride type is being referenced by std (Standard) and sha (Shared). Because the statement of our model requires multiple indices at many points, it is helpful to state explicitly our indexing conventions. Unless stated otherwise, i indexes individuals, j indexes the good whose demand is being affected, and k indexes the price which is being changed. Additionally, we reserve the use of subscripts for indexing individuals while using superscripts to index components of vectors/matrices.

In describing our general conceptual framework above, we derived our results for a population-wide linear latent demand curve and a homogeneous inattention parameter. In our empirical setting, however, we must account for heterogeneity and the fact that we are working with discrete choice data. We therefore assume throughout that individuals have quasi-linear utility, but misperceive prices due to left-digit bias.

Assumption 3. Each individual i can be described by a vector $\theta_i^{std}, \theta_i^{sha}, V_i^{std}, V_i^{sha}$. Given prices P^{std}, P^{sha}, individual i accepts Standard if and only if

$$V_i^{std} - \$p^{std} - (1-\theta_i^{std})¢p^{std} > 0, \text{ and}$$
$$V_i^{std} - \$p^{std} - (1-\theta_i^{std})¢p^{std} \geq V_i^{sha} - p\$^{sha} - (1-\theta_i^{sha})¢p^{sha}$$

Similarly, individual i accepts Shared if and only if

$$V_i^{sha} - \$p^{sha} - (1-\theta_i^{sha})¢p^{sha} > 0, \text{ and}$$
$$V_i^{sha} - \$p^{sha} - (1-\theta_i^{sha})¢p^{sha} > V_i^{std} - \$p^{std} - (1-\theta_i^{std})¢p^{std}$$

We then define the true demand curves, $Y_i(p) = (Y_i^{std}(p), Y_i^{sha}(p))'$, as the vector of potential outcomes. Specifically, let $Y_i^{std}(p)$, $Y_i^{sha}(p)$ be dummy variables, respectively, taking values 1 if and only if Standard or Shared are chosen when prices were exogenously set to p.

As in our general conceptual framework, it is convenient to introduce a notion of latent demand. Specifically, let

$$Y_i^{*,std}(p^{std}, p^{sha}) = \mathbb{1}\{V_i^{std} - p^{std} = \max\{V_i^{std} - p^{std}, V_i^{sha} - p^{sha}, 0\}\}$$
$$Y_i^{*,sha}(p^{std}, p^{sha}) = \mathbb{1}\{V_i^{sha} - p^{sha} = \max\{V_i^{std} - p^{std}, V_i^{sha} - p^{sha}, 0\}\}$$

It is also convenient to define the notation $\hat{P}_i = (\hat{P}_i^{std}, \hat{P}_i^{sha})'$ where $\hat{P}_i^k = \$P_i^k + (1-\theta_i^k)¢P_i^k$ for $k = std, sha$. Under Assumption 3, we have $Y_i^j(P_i) = Y_i^{*,j}(\hat{P}_i)$ for $j = std, sha$.

Prices at Lyft are determined in an algorithmic way, and we know both the details of the algorithm (at least during our sample period) and can directly observe the inputs to the algorithm. We formalize this institutional knowledge with the following assumption.

Assumption 4. Prices are determined by a formula Prices are determined by a formula

$$\begin{bmatrix} P_i^{std} \\ P_i^{sha} \end{bmatrix} = F(M_i, S_i, U_i)$$

where F is a known function, M_i is the (observed) treatment status from the ongoing pricing experiments, S_i is an indicator for whether or not the observation is in the ongoing pricing experiment, and U_i are the (observed) determinants of price that are not explicitly randomized. Moreover,

$$(M_i, S_i) \perp U_i, \theta_i^{std}, \theta_i^{sha}, V_i^{std}, V_i^{sha}$$

Our discrete choice model of demand implies that individual demand curves cannot possibly be linear unless they are completely flat. We assume throughout, however, that there is some level of aggregation at which *aggregate* demand curves can be thought to be (locally) linear. Specifically, given the institutional details at Lyft, it is convenient to condition on the set of inputs to the pricing algorithm *excluding* the exogenous components of price. It is also convenient to condition on θ_i to allow us to describe how the interpretation of our estimates change if we allow heterogeneity in the inattention parameter.

Assumption 5. Let $\theta_i = (\theta_i^{std}, \theta_i^{sha})'$. Then, aggregate latent demand curves satisfy

$$\begin{aligned}\mathbb{E}[Y_i^{*,std}(p^{std}, p^{sha}) \mid U_i = u, \theta_i = t] &= \min\{\max\{\alpha_{std}(u), t + \beta^{std,std}(u,t)p^{std} \\ &\quad + \beta^{std,sha}(u,t)p^{sha}, 0\}, 1\} \\ \mathbb{E}[Y_i^{*,sha}(p^{std}, p^{sha}) \mid U_i = u, \theta_i = t] &= \min\{\max\{\alpha_{sha}(u,t) + \beta^{sha,sha}(u,t)p^{sha} \\ &\quad + \beta^{sha,std}(u,t)p^{std}, 0\}, 1\}\end{aligned}$$

Moreover, the support of the observed pricing variation is such that the min and max above bind with probability 0.

Remark 1. This assumption requires that conditional on U_i, θ_i, the joint density of V_i^{std}, V_i^{sha} does not vary on its support and that the demand vector is strictly on the interior of the probability simplex for all observed prices. Provided the densities of V_i^{std}, V_i^{sha} do not vary too considerably in a neighborhood of the empirically relevant prices, we can alternatively think of this assumption as a local approximation. In practice, when Lyft estimates aggregate demand curves, they appear linear, so this choice represents a reasonable approximation. As a matter of theory, we expect the own-price β terms to be negative and the cross-price β terms to be positive. We will not impose this restriction ex ante but these basic predictions appear to be strongly supported by the data.

Remark 2. Armstrong and Vickers (2015) show that in quasi-linear models such as the one considered here, cross-price derivatives must be symmetric. Given Assumption 5, this implies that $\beta^{std,sha}(u,t) = \beta^{sha,std}(u,t)$. As in the previous remark, we make no ex ante restrictions on our estimators to ensure such conformity, although as we will see, this prediction appears to hold in our data as well.

Proposition 5. *Suppose the data are generated according to Assumptions 3–5. Let p^{std}, p^{sha}, $\$^{std}$, $\$^{sha}$ be integers such that the* min *and* max *in Assumption 5 are non-binding. Then*

$$\begin{aligned}&\lim_{p\to(\$^{std})^+} \mathbb{E}[Y_i^j(p, p^{sha}) \mid U_i = u, \theta_i = t] \\ &\quad - \lim_{p\to(\$^{std})^-} \mathbb{E}[Y_i^j(p, p^{sha}) \mid U_i = u, \theta_i = t] = \beta^{j,std}(u,t)t^{std} \\ &\lim_{p\to(\$^{sha})^+} \mathbb{E}[Y_i^j(p^{std}, p) \mid U_i = u, \theta_i = t] \\ &\quad - \lim_{p\to(\$^{sha})^-} \mathbb{E}[Y_i^j(p^{std}, p) \mid U_i = u, \theta_i = t] = \beta^{j,sha}(u,t)t^{sha}\end{aligned}$$

This proposition shows that conditional on the inputs to the pricing algorithm $U_i = u$ and fixing a value of the inattention parameter $\theta_i = t$, the size of the discontinuity when price $k = std, sha$ is varied is exactly t^k times the slopes of the demand curve with respect to price k. This suggests the following empirical strategy for estimating an inattention parameter: fix the value of U_i and measure the size of the discontinuity when crossing an integer price as well as own/cross-price slopes. Then, the ratio between the discontinuity and the slope identifies inattention. Proposition 5 conditions on the inattention parameter θ_i and therefore does not directly justify the above strategy. Nonetheless, the following corollary helps clarify what we precisely learn if we implement the above strategy.

Corollary 2. *Suppose the data are generated according to Assumptions 3–5. Let* p^{std}, p^{sha}, $\$^{std}$, $\$^{sha}$ *be integers such that the* min *and* max *in Assumption 5 are always non-binding. Then*

$$\lim_{p\to(\$^{std})^+}\mathbb{E}[Y_i^j(p,p^{sha})\,|\,U_i=u]-\lim_{p\to(\$^{std})^-}\mathbb{E}[Y_i^j(p,p^{sha})\,|\,U_i=u]$$
$$=\mathbb{E}[\beta^{j,std}(u,\theta_i)\,|\,U_i=u]\mathbb{E}[W_i^{j,std}\theta_i^{std}\,|\,U_i=u]$$
$$\lim_{p\to(\$^{sha})^+}\mathbb{E}[Y_i^j(p^{std},p)\,|\,U_i=u]-\lim_{p\to(\$^{sha})^-}\mathbb{E}[Y_i^j(p^{std},p)\,|\,U_i=u]$$
$$=\mathbb{E}[\beta^{j,sha}(u,\theta_i)\,|\,U_i=u]\mathbb{E}[W_i^{j,sha}\theta_i^{std}\,|\,U_i=u]$$

where

$$W_i^{j,k}=\frac{\beta^{j,k}(U_i,\theta_i)}{\mathbb{E}[\beta^{j,k}(U_i,\theta_i)\,|\,U_i]}$$

for $j=std,sha$ *and* $k=std,sha$.

This proposition implies that if we fix a value of U_i and divide the demand discontinuity at an integral dollar amount by slope of the average latent demand curve, then we will identify a *weighted* average inattention parameter. In particular, values of θ_i where a higher density of individuals are on the margin between switching their purchase decision will receive more weight. Moreover, assuming that the own-price slopes are all non-positive and the cross-price slopes are all non-negative, the weights will always be positive. This is exactly analogous to the usual LATE interpretation of IV as noted by Angrist and Imbens (1994). We therefore refer to the inattention parameters identified via this strategy as the local average inattention.

Finally, consider the following analog to Corollary 1.

Corollary 3. *Consider the setting of Corollary 2. Suppose that* $\mathrm{Cov}(W_i^{std,std},\theta_i^{std}|U_i=u)=\mathrm{Cov}(W_i^{sha,std},\theta_i^{std}|U_i=u)=0$. *Then for all* u,

$$\frac{\lim_{p\to(\$^{std})^+}\mathbb{E}[Y_i^{std}(p,p^{sha})\,|\,U_i=u]-\lim_{p\to(\$^{std})^-}\mathbb{E}[Y_i^{std}(p,p^{sha})\,|\,U_i=u]}{\mathbb{E}[\beta^{std,std}(u,\theta_i)\,|\,U_i=u]}$$
$$=\frac{\lim_{p\to(\$^{std})^+}\mathbb{E}[Y_i^{sha}(p,p^{sha})\,|\,U_i=u]-\lim_{p\to(\$^{std})^-}\mathbb{E}[Y_i^{sha}(p,p^{sha})\,|\,U_i=u]}{\mathbb{E}[\beta^{sha,std}(u,\theta_i)\,|\,U_i=u]}$$

Similarly, if $\mathrm{Cov}(W_i^{sha,sha},\theta_i^{sha}\,|\,U_i=u)=\mathrm{Cov}(W_i^{std,sha},\theta_i^{sha}\,|\,U_i=u)$, *then*

$$\frac{\lim_{p\to(\$^{sha})^+}\mathbb{E}[Y_i^{sha}(p^{std},p)\,|\,U_i=u]-\lim_{p\to(\$^{sha})^-}\mathbb{E}[Y_i^{sha}(p^{std},p)\,|\,U_i=u]}{\mathbb{E}[\beta^{sha,sha}(u,\theta_i)\,|\,U_i=u]}$$
$$=\frac{\lim_{p\to(\$^{sha})^+}\mathbb{E}[Y_i^{std}(p^{std},p)\,|\,U_i=u]-\lim_{p\to(\$^{sha})^-}\mathbb{E}[Y_i^{std}(p^{std},p)\,|\,U_i=u]}{\mathbb{E}[\beta^{std,sha}(u,\theta_i)\,|\,U_i=u]}$$

Compared to Corollary 1, Corollary 3 implies a somewhat weaker test of our left-digit bias model. Intuitively, it states that provided heterogeneity in inattention is not correlated with the slope of the latent demand curve, we should expect that the inattention parameter estimated from cross-price elasticities should be the same as the inattention estimated from own-price elasticities. The comparison thus constitutes a joint test of the inattention model *and* the assumption that conditional on the inattention model being true, heterogeneity is "mild". In our empirical results, we find that the predictions of Corollary 3 hold up remarkably well, which we tentatively interpret as evidence that not only is the left-digit bias model a reasonable description of the data, but that heterogeneity in the inattention parameter is relatively mild.

When we estimate inattention parameters in the results to follow, we will always be implementing some variant of Corollary 2. However, as will be seen, the available variation in the respective settings as well as computational constraints will force us to implement different estimators in our observational setting compared to our experimental setting.

4. OBSERVATIONAL EVIDENCE

4.1. *Graphical analysis*

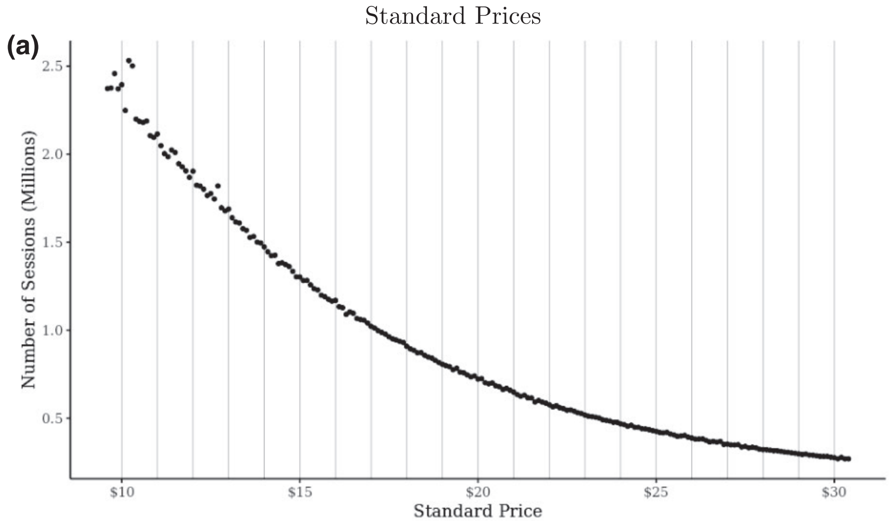
We begin by analyzing observational data using the ∼600 million sessions that occurred between February and August of 2019 where Standard rides or Standard and Shared rides were offered. Figure 2 plots the Standard and Shared price distributions for all rides with price offers between \$8.50 and \$30.50.[7] Each dot in the graph illustrates the number of sessions with prices in a 10-cent range (\$14.60–\$14.69, \$14.70–\$14.79, etc.). Consistent with our knowledge that Lyft's pricing algorithm during this time period was agnostic to round dollar marks, the price distributions are very smooth.[8] Importantly, there is no evidence of extra mass right below dollar amounts (*e.g.* there are a similar number of prices offered between \$12.90 and \$12.99 as \$13.00 and \$13.09). Appendix Figure A2 examines 1-cent bins and also does not find evidence of bunching below dollar amounts (*e.g.* the sessions with a price of \$12.99 are similar to those with a price of \$13.00).

Given the smoothness in the price distributions, we next show graphically how passengers respond to prices that are just above versus just below dollar values. The outcome of interest is whether a conversion occurred (the passenger accepted one of the prices being offered). Panel A of Figure 3 shows the fraction of sessions that resulted in a conversion (of either product) for different Standard prices. Overall, the probability of not converting increases as the Standard price increases. However, consistent with Proposition 1, there are clear discontinuities in non-conversion occurring at every dollar value. Prices that are just over a dollar (*e.g.* \$17.00–\$17.09) are discontinuously more likely to have non-conversion than prices just under a dollar (*e.g.* \$16.90–\$16.99). Visually, crossing over a dollar threshold discontinuously increases the non-conversion rate by approximately 1 percentage point at each dollar value.

Panel B of Figure 3 shows the fraction of sessions that result in a conversion of the Standard product for different Standard prices. Once again, we see large discontinuities at dollar values such that Standard prices just under a dollar threshold are much more likely to lead to a Standard conversion than those that are just above a dollar threshold. Lastly, Panel C plots the fraction of

7. We restrict the x-axis to this dollar range since most prices are in this range (66%) and to make it visually easier for the reader. The full distributions, ranging from \$3 to \$100 can be seen in Appendix Figure A1.

8. An anonymous referee pointed out to us that separately plotting only the *marginal* distributions of Standard and Shared prices may mask important richness in their relationship. For instance, one might be concerned that if Shared price were a fixed percent discount off of Standard price, then the cents amount for Shared is a deterministic function of Standard. As a result, jumps at integer prices may at some points simultaneously occur for both Standard and Shared, thus muddying the interpretation of the discontinuity. While we have some restrictions on how much detail we can provide on the relationship between shared and standard prices, based on internal conversations at Lyft, we have reason to believe that these concerns are minor. The relationship between Standard and Shared prices is quite variable, with the interquartile range of Shared price as a fraction of Standard price being roughly 20%, almost the same size as the average discount. Nonetheless, in Appendix B, we present additional evidence to address these concerns by showing that the same qualitative results continue to hold if we restrict ourselves to sessions where only a Standard price was present. In addition, within our experimental dataset, concerns that too many factors are changing at once are no longer present. We use only variation in cents amount that our experiment explicitly generated, and because we are the designers of the experiment, we know exactly how the various treatments affected prices. Reassuringly, the quantitative results in this section and the next section where we analyze the experimental data are in agreement.



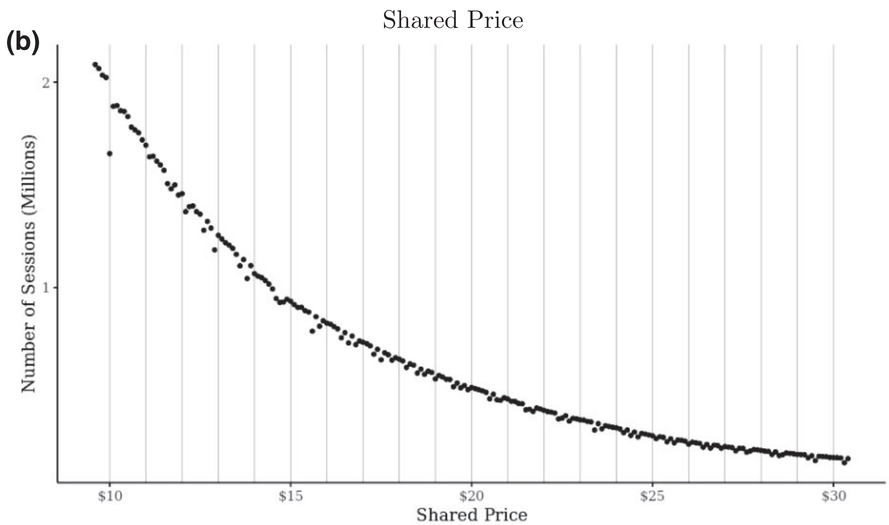


FIGURE 2

Distribution of prices

Notes: This figure plots the distribution of Standard and Shared prices. The sample used to construct this figure is all Lyft pricing sessions with Standard and Shared prices between \$9.50 and \$30.05 from February 2019 through August 2019 inclusive. We additionally exclude sessions that contained prices for Lyft's Shared Saver product, because the way these sessions were priced makes the subsequent analysis difficult to interpret. Each point above corresponds to a price truncated after the first decimal (so, for example, a price of \$15.07 becomes \$15.00). Panel (a) plots the number of sessions for each truncated Standard price while Panel (b) plots the number of sessions for each truncated Shared price. The vertical lines mark integer prices.

sessions that end with a conversion of the Shared product for different Standard prices. Consistent with substitution occurring, when the Standard price is just over a dollar value, passengers are more likely to take the Shared ride as opposed to when the Standard price is just under the dollar threshold.

A natural question is why the Standard conversion rate is upward sloping between dollar values in Panel B of Figure 3. The reason for this is because we are not examining a particular Lyft trip with a randomized price offering between \$10 and \$30. Instead, we are observing

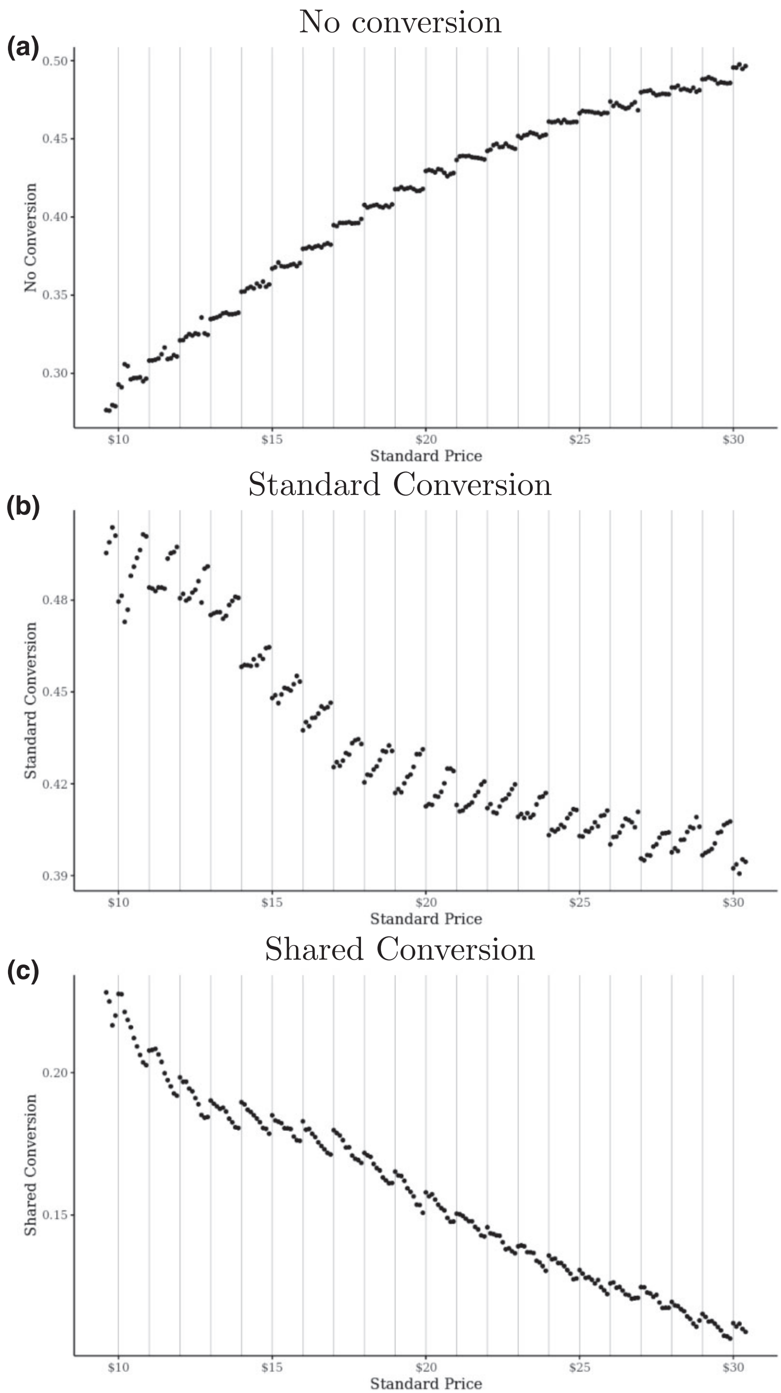


FIGURE 3

Conversion rates by standard price

Notes: This figure plots conversion outcomes across different Standard prices truncated after the first decimal. Our sample is constructed as in Figure 2, and each point again corresponds to a price truncated to the first decimal. Panel (a) plots the proportion of sessions for each truncated Standard price that did not covert. Panel (b) plots the proportion of sessions for each truncated Standard price that converted to Standard. Panel (c) plots the proportion of sessions for each truncated Standard price that converted to Shared. The vertical lines mark integer prices.

varying prices alongside the types of trips varying (*e.g.* longer trips = higher prices). The fact that the good/service itself is changing with price is causing the upward slope between dollar values. The intuition for this result is as follows. Assume, for example, that app users have complete left-digit bias (they only look at the dollar value and ignore cents completely). Then, as the standard price moves within a dollar range (*e.g.* from $14.00 to $14.99), the price does not seem to be changing at all (because of full left-digit bias), but it feels like a better deal because the $14.99 ride is probably a bit longer than the $14.00 ride and so the price being offered feels like a good deal. Thus, with complete left-digit bias, the $14.99 feels like a bargain compared to the $14.00 ride. This leads to an upward slope. An additional contributing factor is that as the standard price rises from $14.00 to $14.99, it starts to look more attractive relative to the shared price which likely crossed over a dollar value at some point during that range (and likely looks more attractive compared to other substitutes as well). As can be seen in Appendix B, the upward slope within a dollar value range is not nearly as strong when restricting the sample to markets with only a standard ride option.

Figure 4 provides analogous graphs to those in Figure 3, but does so for conversion at different Shared prices. Relative to being just under a dollar value, when Shared prices are just over a dollar value, non-conversions go up, conversions of the Shared product go down, and conversions of the Standard product go up. These results qualitatively match those shown in Figure 3 when focusing on Standard prices.

Figures 3 and 4 group prices into 10-cent bins to show evidence of discontinuities in demand at dollar values across the pricing distribution. Binning prices in this way is convenient for showing the full results, but introduces ambiguity regarding the exact point at which demand drops or rises. In particular, one could imagine that rather than demand dropping discontinuously as the price increases from a 99-cent mark to the dollar value above, it could be that the slope of the demand curve becomes “steep” around dollar values. Some retail outlets, for example, set prices that end with 95 or 97 cents rather than 99 cents to potentially prevent customers from rounding up prices in their head. Inattention mechanisms for left-digit bias on the other hand typically suggest that consumers fail to even process or consider non-left-digits and thus there should be a discontinuity rather than a steepening in demand around dollar thresholds.

In Figure 5, we plot the change in conversion for Standard prices at each cent amount within 15 cents of a round dollar value. We stack the data such that each dot represents the weighted average conversion for prices that end in each cent value around dollar amounts from $10 to $30. All three panels show a clear discontinuity that is occurring right at dollar thresholds, as opposed to a steepening behavioral response to prices as they approach dollar thresholds.[9] Overall, the graphical evidence is compelling and consistent with the predictions of left-digit bias presented in Section 2.

4.2. *Quantitative analysis*

We complement the graphical analysis with quantitative methods that more precisely estimate the level of inattention to the cent amounts in Lyft prices. Heuristically, the discussion in Section 3.2 shows that we need to compare the effect of a large price change (*e.g.* the change in quantity demanded due to a price change from $15.40 to $14.40) to the effect of the price crossing an integer number (*e.g.* the change in quantity demanded due to a price change from $15.00 to $14.99).

9. We show just the Standard price graphs to conserve space, but also because at the cent-level, Shared prices are not smooth for idiosyncratic institutional reasons—as discussed in association with Appendix Figure A1.

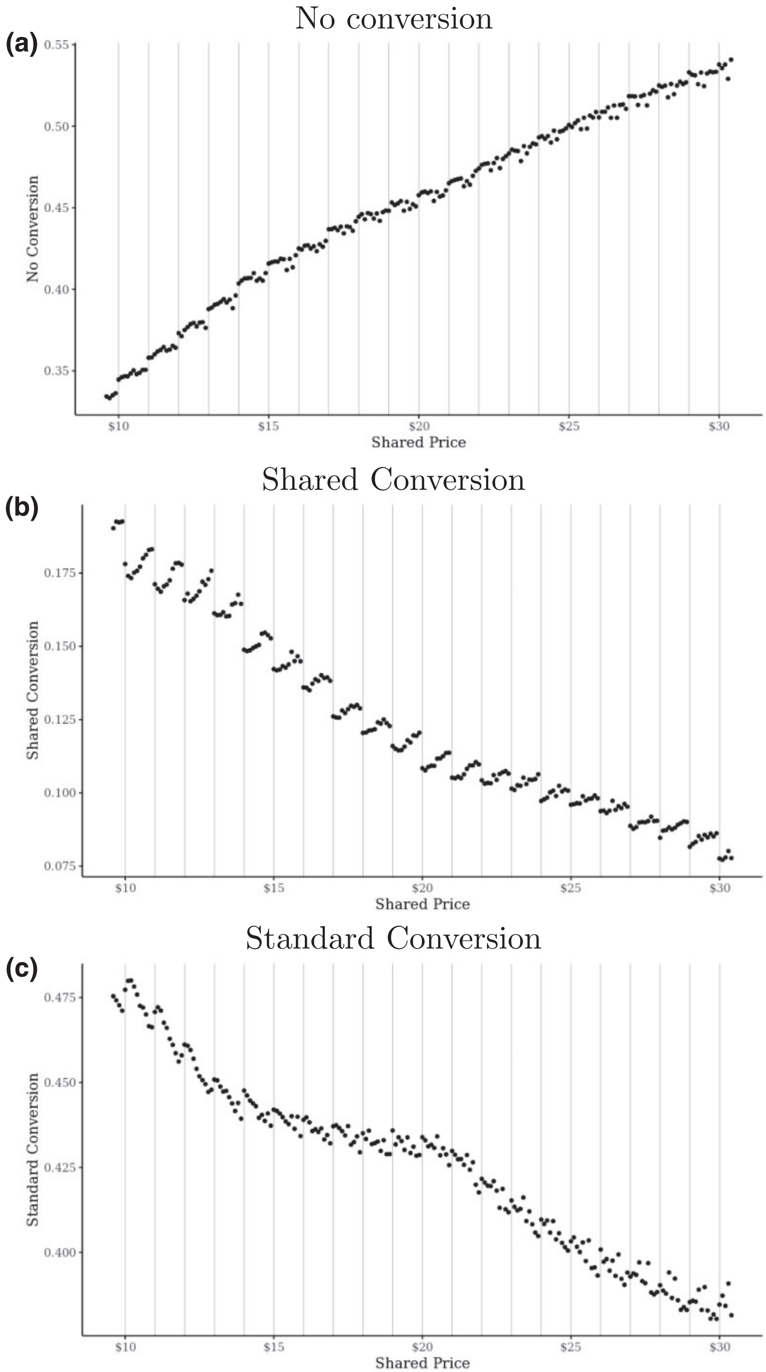


FIGURE 4
Conversion rates by shared price

Notes: This figure plots conversion outcomes across different Shared prices truncated after the first decimal. Our sample is constructed as in Figure 2, and each point again corresponds to a price truncated to the first decimal. Panel (a) plots the proportion of sessions for each truncated Shared price that did not covert. Panel (b) plots the proportion of sessions for each truncated Shared price that converted to Shared. Panel (c) plots the proportion of sessions for each truncated Shared price that converted to Standard. The vertical lines mark integer prices.

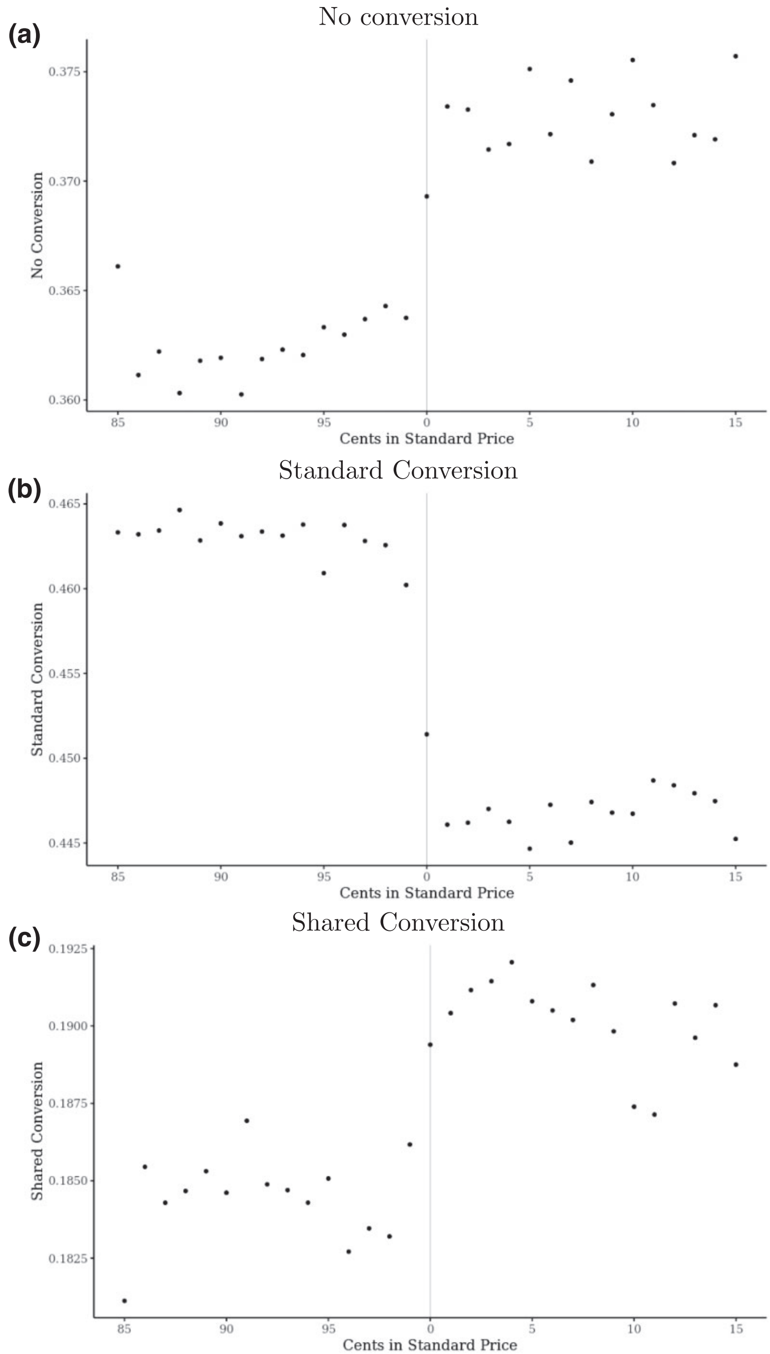


FIGURE 5
Conversion rates by cents in standard price

Notes: This figure plots conversion outcomes across different cent amount in Standard prices. Our sample is constructed as in Figure 2, but we subset to prices with cent amount in price ranging from 85 to 99 and from 0 to 15. Each point corresponds to a cent amount. Panel (a) plots the proportion of sessions for each cent amount that did not covert. Panel (b) plots the proportion of sessions for each cent amount that converted to Standard. Panel (c) plots the proportion of sessions for each cent amount that converted to Shared. The vertical line marks where the cent amount is exactly 0 (*i.e.* the price is integer).

Throughout Section 3.2, we assumed for mathematical convenience that the latent demand curves were linear in price, so the relevant measure of price sensitivity was the slope/derivative. Inattention was then identified in terms of the ratio between the slope of the demand curve and the size of the discontinuity. In this section, we implement this idea as closely as possible with our observational data.

We begin by estimating the own- and cross-price slopes for different baseline price-points. Before discussing our empirical implementation, it is helpful to discuss how one would ideally identify these slopes. In the best-case scenario, Lyft would run an experiment where, for each price, it would randomly serve consumers either a $\$X$ price increase over the baseline or a $\$X$ decrease in price. In that case, the slope of the demand curve could be estimated simply by dividing the treatment effect in this experiment by the difference in price between these two treatments, *i.e.* $\$2X$. For each integral dollar amount, we would then, as in Corollary 2, divide the size of the discontinuity around this integer price by the estimated slope of the demand curve for baseline prices near this integer price and interpret this ratio as an estimate of inattention.

The experimental data we do have access to it exogenous, as required, but it is not quite as clean as the hypothetical experiment described above. Specifically, as discussed in Section 2, the true mapping between experimental treatments and realized price changes is determined by a complex algorithm mapping experimental treatments to price changes. Loosely speaking, this algorithm works by first drawing an exogenous "multiplier" and serving a final price to the consumer roughly equal to the baseline price times this multiplier. In practice, however, Lyft's internal pricing algorithm applies a complex set of additional transformations that, depending on the setting at hand, can obscure the exact relationship between the multiplier and the price change that actually gets implemented. As a result, rather than estimating a straightforward treatment effect as in our hypothetical ideal experiment described in the previous paragraph, we instead adopt an instrumental variables strategy.

Specifically, for each price (Standard and Shared), we group sessions by the dollar amount closest to the price produced by Lyft's pricing algorithm before experimental variation is added, which we call counterfactual prices. We then estimate the (2×2) matrix of own- and cross-price elasticities for the Standard and Shared products by using the exogenously varied multipliers as instruments for price. Formally, we estimate elasticities using 2SLS with first stage

$$\begin{aligned} P_{std} &= \gamma_{1,0} + \gamma_{1,1} M_{std} + \gamma_{1,2} M_{sha} + \nu_{std} \\ P_{sha} &= \gamma_{2,0} + \gamma_{2,1} M_{std} + \gamma_{2,2} M_{sha} + \nu_{sha} \end{aligned} \tag{6}$$

and second stage

$$\begin{aligned} Y_{std} &= \alpha_{1,0} + \beta_{1,1} P_{std} + \beta_{1,2} P_{sha} + \varepsilon_{std} \\ Y_{sha} &= \alpha_{2,0} + \beta_{2,1} P_{std} + \beta_{2,2} P_{sha} + \varepsilon_{sha} \end{aligned} \tag{7}$$

where the subscripts std, sha, respectively, index Standard and Shared, p represents price, y is an indicator for whether or not the user in a given session converted to either Shared or Standard, and m represents the exogenous multipliers.

The β's in equation (7) correspond to own- and cross-price slopes of demand. Because the price variation induced by the random multipliers occurs for counterfactual prices with all different cent endings, we can interpret these slopes as being approximately equal to the slopes from the price change of exactly one dollar as described in Proposition 4.[10]

10. We provide a more rigorous justification for this claim in Appendix C.

The own- and cross-price slopes that we calculate for each dollar amount are shown in Panels A and B in Figures 6 and 7. Panel A of Figure 6 illustrates an own-price slope for Standard price that ranges from −0.02 to −0.05 with lower-priced sessions measured as the most inelastic. Panel B of Figure 6 shows cross-price slops for Standard price ranging from 0.005 to 0.02. Analogously, Panel A of Figure 7 shows an own-price slope for Shared prices ranging from −0.01 to 0.03 and Panel B shows a cross-price slope for Shared prices ranging from 0.004 to 0.016. The own- and cross-price slopes that we estimate are in line with estimates that have been calculated by other research teams using Lyft and Uber data.[11]

We next estimate own- and cross-price effects when a price crosses over a dollar threshold by comparing the conversion rates for prices just above versus just below the round dollar amounts (*i.e.* estimating the discontinuities in Panels B and C of Figures 3 and 4). We plot the size of the own- and cross-price discontinuities by measuring the size of the discontinuity for price changes just above and below dollar values in Panels C and D of Figures 6 and 7. The own- and cross-price discontinuities are roughly half the size of the slopes shown in Panels A and B of the same figures. Interestingly, the slopes of the lines in panels A and C (and also panels B and D) are moving in the same direction. For example, the overall own-price slope is larger in magnitude for lower-priced sessions, just as the discontinuous jumps are larger for lower-priced sessions. Another way of summarizing this data pattern is that for rides where passengers are highly price-sensitive, they are also relatively more sensitive to changes in the left-most digit.

The inattention model described in Section 3.2 allows us to make sense of the patterns we see in the data. Appealing to a regression discontinuity argument, we expect that the size of the discontinuity observed at each integer price should largely be a causal quantity, and our slopes are estimated from experimental variation. Our setting therefore closely resembles the setting of Corollary 2, which allows us to interpret the fact that the discontinuity is roughly half of the slope as corresponding to an inattention parameter of 0.5.

While our empirical strategies attempt to follow the theoretical derivations as closely as possible, there are a number of subtle distinctions between Corollary 2 and our estimation approach. First, Corollary 2 conditions on all of the details of the pricing algorithm, U_i, while we only condition on the counterfactual price via our subsetting. A second, related concern is that as a result of this imperfect conditioning, our 2SLS strategy actually identifies a weighted average slope that may be difficult to interpret (see, for instance, Angrist, Graddy, and Imbens (2000)). In principle, this weighting could invalidate the comparisons we are making since it means that the slopes corresponding to crossing dollar thresholds could be estimated from a different population than the populations used in each regression discontinuity. With unlimited computational resources, it is possible to correct for this weighting issue, which we do on the much smaller experimental sample in Section 5.3. We adopt the IV approach described above in this section because 2SLS can be easily implemented in SQL code, which allows us to run our regressions on datasets with hundreds of millions of observations. Reassuringly, the results in Section 5.3 produce quantitatively similar results, suggesting that the distinctions raised here are unlikely to be substantially driving our results.

11. The typical Standard slope reported in Figure 6 is around −0.035, which, given the average Standard price of \$16.89 and the average Standard conversion of 0.64 translates to an elasticity of roughly 0.92. This is somewhat higher than what is reported in Cohen *et al.* (2016), whose elasticities range from −0.46 to −0.66. There are differences between our setting and the one studied by Cohen *et al.* (2016) that might explain these elasticity differences. For example, Lyft had a shared ride option for many markets in our data, which represented a fairly close substitute for the Standard product. Also, the variation in prices that we use occurs randomly across all sessions while Cohen et al. exploit variation in real-time pricing that is triggered only during undersupplied market conditions.

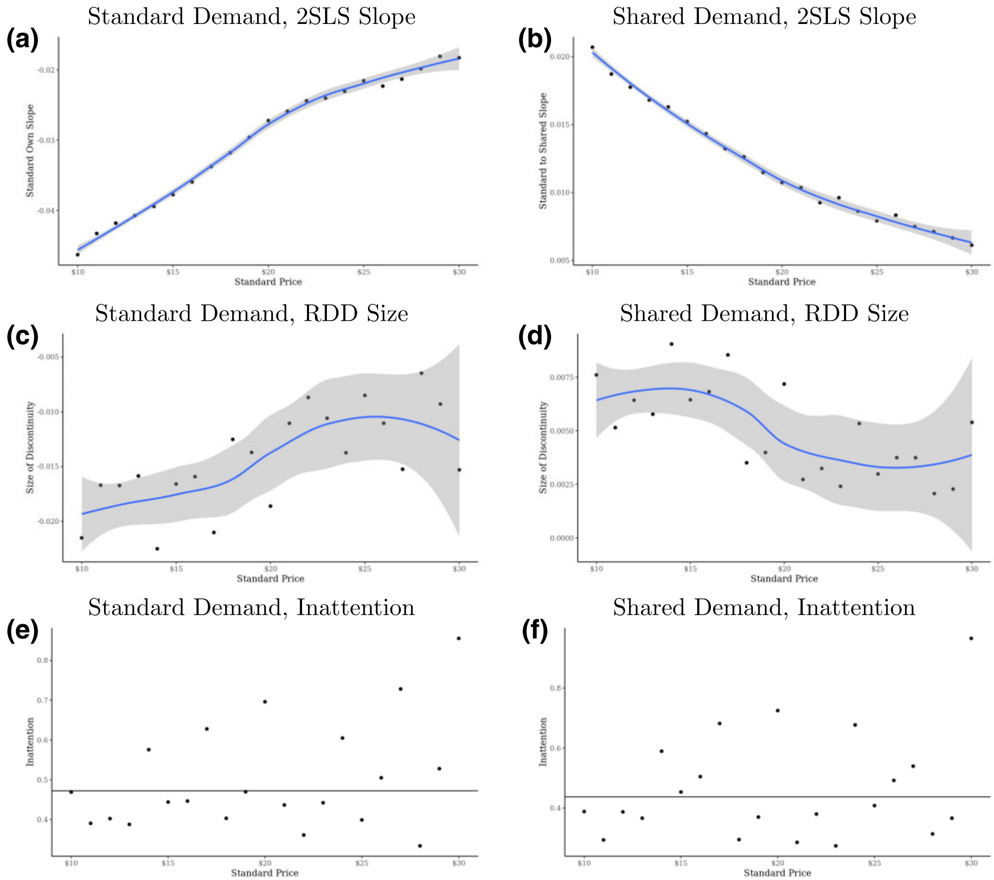


FIGURE 6

Slope at each dollar: standard price

Notes: This figure shows the steps described in the text for estimating an inattention parameter with respect to Standard prices. The sample for these plots is identical to the sample for the previous plots, except that for our 2SLS estimates, we restrict our attention to the 10% of sessions that had exogenous price variation. Each point corresponds to a fixed dollar amount of Standard price. The column on the left uses Standard conversion as the outcome variable. The column on the right uses Shared conversion as the outcome variable. In the top row, each point is constructed by grouping together observations whose counterfactual (*i.e.* in the absence of the exogenous variation) Standard price gets rounded to the same integer value and computing a slope for each group. The middle row computes the elasticity of Standard/Shared corresponding to the discontinuities observed at each dollar amount of Standard price (where we take the change in price to be the one cent drop from $\$X.00$ to $\$(X-1).99$). The size of the discontinuity is computed by taking the difference in conversion for sessions where Standard prices were between \$X.00 and \$X.09 compared to sessions where Standard prices were between \$(X-1).90 and \$(X-1).99. Finally, the bottom row estimates an inattention parameter by taking the ratio of the elasticity in the middle row and the elasticity of the top row times 0.01. The horizontal line in the bottom row is the weighted average of the points, where the weights are proportional to the number of sessions with a given price.

A third distinction between our empirical strategy and our theoretical framework is that the "slope" in Corollary 2 is the slope of the latent demand curve whereas we simply use the slope of the demand curve, measured for a sufficiently large change in price. In principle, we could have alternatively tried to more closely estimate the slope of the latent demand curve by jointly estimating inattention and latent demand curve slope by explicitly modifying the structural equation in equation (7) to explicitly depend on θ. We would then estimate θ and the demand curve slopes via a nonlinear instrumental variables strategy. Again, our primary motivation for not doing this

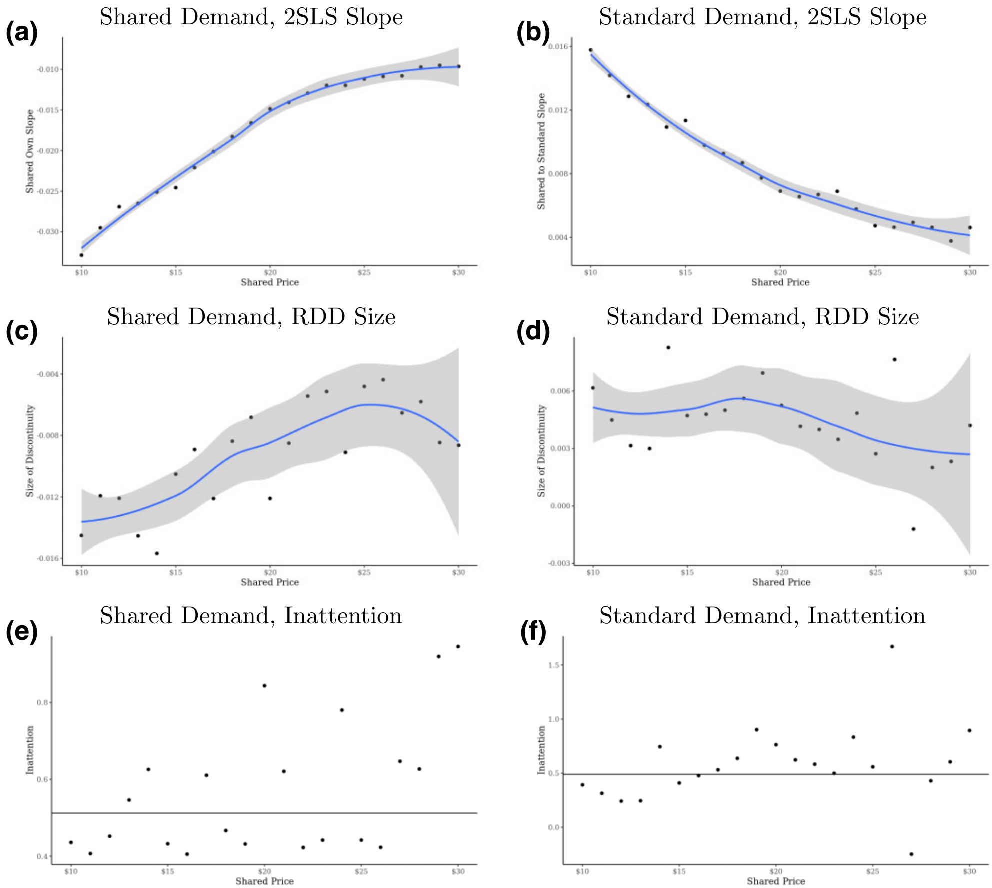


FIGURE 7

Slope at each dollar: shared price

Notes: This figure shows the steps described in the text for estimating an inattention parameter with respect to Shared prices. The sample for these plots is identical to the sample for the previous plots, except that for our 2SLS estimates, we restrict our attention to the 10% of sessions that had exogenous price variation. Each point corresponds to a fixed dollar amount of Shared price. The column on the left uses Shared conversion as the outcome variable. The column on the right uses Standard conversion as the outcome variable. In the top row, each point is constructed by grouping together observations whose counterfactual (*i.e.* in the absence of the exogenous variation) Shared price gets rounded to the same integer value and computing an elasticity for each group. The middle row computes the elasticity of Shared/Standard corresponding to the discontinuities observed at each dollar amount of Shared price (where we take the change in price to be the one cent drop from $\$X.00$ to $\$(X-1).99$). The size of the discontinuity is computed by taking the difference in conversion for sessions where Shared prices were between \$X.00 and \$X.09 compared to sessions where Shared prices were between \$(X-1).90 and \$(X-1).99. Finally, the bottom row estimates an inattention parameter by taking the ratio of the elasticity in the middle row and the elasticity of the top row times 0.01. The horizontal line in the bottom row is the weighted average of the points, where the weights are proportional to the number of sessions with a given price.

is computational because we needed to accommodate the sheer size of our data. In Appendix B, we show that this estimation choice is also unlikely to have significantly affected our results.

We plot these inattention parameters for each dollar value in Panels E and F in Figures 6 and 7. The average inattention across dollar values is approximately 0.5, but there is some variation. For example, inattention is especially high at the \$20 and \$30 values which could be consistent with certain models of left-digit bias (as was discussed in Section 2).

The local average inattention parameters using both own- and cross-price slopes are shown in Table 2. The inattention parameter varies between 0.43 and 0.51, regardless of whether we

TABLE 2
Inattention parameter from observational data

	Standard	Shared
Standard	0.47	0.44
	(0.02)	(0.03)
Shared	0.49	0.51
	(0.06)	(0.03)

Notes: This table shows estimates for inattention parameters using the observational data. Each entry is a single estimate of inattention, identified off of a single own/cross-price elasticity. The rows correspond to which price is being varied while the columns correspond to the product whose demand we are looking at (so, for example, the top right number corresponds to inattention identified by the cross-price elasticity of varying Standard price and looking at the effect on Shared demand). The inattention parameters are the same as those implied by Panels E and F of Figures 6 and 7. The point estimates in this table are computed by taking the weighted average of the inattention estimates amounts in these figures (one estimate for each dollar amount), with weights proportional to the number of sessions with a given price (rounded to the nearest dollar). The standard errors are computed using the weighted standard error formula corresponding to this weighting.

focus on the own- or cross-price effects of the Standard or Shared product. We can interpret this level of inattention in a few different ways. First, an inattention parameter of 0.5 means that on average, people perceive the actual price to be the dollar amount of the price plus 50% of the cent amount. Another way to interpret this estimate is that the one-cent drop from the round dollar amount to the 99-cent price below has the same effect as a 50-cent drop in price to the average user. Put a final way, if we lower prices by a dollar from $\$X.00$ to $\$(X-1).00$ one cent at a time, 50% of the increase in demand comes from the very first cent.

4.3. *Heterogeneity*

The previous section demonstrated remarkable stability in the inattention parameter whether examining Standard or Shared prices as well as own- or cross-price impacts. We can also test whether inattention to non-left-digits varies across different types of individuals and trip types. There are mechanisms that one could posit for why the inattention parameter might vary. For example, passengers who always use the same product (either Standard or Shared) have fewer prices they need to look at and therefore might pay closer attention to the one price on which they are focused. Or, experienced customers may be more attentive to prices because they are more familiar with the app and can therefore use their attention to focus on the prices as opposed to other features of the app.

To examine heterogeneity across passenger types, we split the data by high versus low-frequency users (a high-frequency user is defined as having 10+ sessions in the 28 days prior to the current session; (this approach is similar to tests in List, 2003, 2004), users who show a strong preference for one product versus those who try both products (defined as having >90% of all rides in the 28 days prior being either Standard or Shared rides), and business versus non-business users (defined as having a registered enterprise account). To explore heterogeneity across trip types, we split the data by airport versus non-airport originating trips, weekday versus weekend trips, and day versus night (9PM–6AM) trips. Lastly, we also estimate inattention in sessions where only the Standard product was available compared to in sessions where both the Standard and Shared products are available.

We identify inattention in the same manner as described in the previous section (computing the ratio of own-price slope around dollar values to the overall price slope), and we compute the inattention parameter separately for Standard and Shared prices. The first two columns of Table 3 display the inattention parameters with accompanying standard errors.

TABLE 3
Heterogeneity in inattention

	Standard θ	Shared θ	Standard elasticity	Shared elasticity
High frequency	0.51	0.53	−1.01	−1.91
	(0.02)	(0.03)	(0.04)	(0.09)
Low frequency	0.45	0.49	−1.25	−2.51
	(0.02)	(0.03)	(0.04)	(0.09)
One product	0.45	0.50	−0.75	−1.84
	(0.02)	(0.04)	(0.03)	(0.09)
Not one product	0.52	0.54	−2.21	−2.26
	(0.03)	(0.03)	(0.07)	(0.09)
Had business account	0.45	0.48	−0.76	−2.47
	(0.04)	(0.04)	(0.02)	(0.09)
No enterprise account	0.47	0.51	−1.20	−2.25
	(0.03)	(0.03)	(0.05)	(0.10)
Airport	0.48	0.49	−0.95	−2.23
	(0.04)	(0.05)	(0.02)	(0.07)
Not airport	0.48	0.51	−1.18	−2.26
	(0.02)	(0.03)	(0.05)	(0.10)
Weekday	0.45	0.52	−1.14	−2.26
	(0.02)	(0.03)	(0.04)	(0.10)
Weekend	0.52	0.50	−1.21	−2.23
	(0.02)	(0.03)	(0.05)	(0.10)
6AM–9PM	0.47	0.52	−1.18	−2.29
	(0.02)	(0.03)	(0.04)	(0.10)
Not 6AM–9PM	0.50	0.49	−1.11	−2.13
	(0.03)	(0.03)	(0.05)	(0.10)
Standard only	0.39		−0.66	
	(0.03)		(0.03)	
Standard and Shared	0.47		−1.16	
	(0.022)		(0.04)	

Notes: This table shows heterogeneity in estimated inattention parameters and elasticities across a number of heterogeneity cuts. Each pair of rows corresponds to a single heterogeneity cut, and each row labels what subset of the data is used to estimate inattention and elasticity. The definition of "One product" is identical to the definition in Table 1. The point estimates and standard errors are computed analogously to how they are computed in Table 2, using weighted averages of the point estimates for a fixed dollar amount, but subsetting according to the given heterogeneity cut. For elasticity, we aggregate the dollar-wise elasticities computed by 2SLS in Panels (a) and (b) of Figures 6 and 7 in the same way as we aggregate inattention estimates. The first column displays our estimates for inattention to Standard prices. The second column displays our estimates for inattention to Shared prices. The third column displays our estimates of Standard own-price elasticity. The fourth column displays our estimates of Shared own-price elasticity.

We find relatively stable inattention parameter values for the different sub-groupings of the data. For both Standard and Shared prices, we estimate values of inattention ranging from 0.39 to 0.54, with the vast majority of values being very close to 0.5. Of the 13 comparisons, only one of them yields a statistically significant difference at the 5% level (before adjusting for multiple hypothesis testing).[12]

A potential concern with this heterogeneity analysis is that the sub-groupings that we use are not segmenting passengers or trips into meaningfully different groups. However, we can provide one piece of evidence that these sub-groupings are creating meaningfully different passenger

12. The one comparison that is statistically significant is people appear more inattentive in cities where both Standard and Shared products are available compared to cities with just a Standard option. This is consistent with passengers being more attentive when they only have one price to consider. However, there are many other stories that could explain this difference as well.

segments. In Columns 3 and 4 of Table 3, we report the average own-price elasticity for each sub-grouping that we analyze.[13] In contrast to the stability in the inattention parameters, own-price elasticities (especially for the more precisely estimated own-price elasticity of the Standard product) vary meaningfully between groups. For example, passengers who always choose the same product, have a business account, originate their ride from an airport, or live in a city with just the Standard product available, are all significantly more inelastic than their counterparts.

How can elasticities vary so much and yet the inattention parameter remain stable? Consider the following example: passengers who always choose the same product have an elasticity that is 3 times smaller than passengers who frequently switch between using the Standard and Shared products (-0.75 versus -2.21). And yet, they have very similar inattention parameters (0.45 versus 0.52). Intuitively, this means that the slope of the demand curve is much steeper for passengers who always choose the same product, but the discontinuity that occurs at round dollar values is also much larger for this subgroup. Thus, the ratio of the overall elasticity to the elasticity just around dollar values is similar, resulting in equivalent inattention parameters.

The fact that the sub-groups in our analysis have different elasticities but remain with similar levels of inattention is interesting, but in the end still does not conclusively demonstrate that there is no meaningful individual heterogeneity in inattention in this context. Even within these sub-groups, there could be meaningful variation in individual characteristics (age, wealth, etc.) that might correlate with differences in inattention. Conclusive evidence requires estimating individual-level inattention parameters and examining the distribution of parameter values. Unfortunately, our data do not allow for precise individual-level estimates. Because the inattention parameter is a ratio, it requires far more observations for a given individual than is available in our data. Thus, we argue that our results are consistent with a surprising level of stability in inattention parameters across individuals but not conclusive.

5. FIELD EXPERIMENTAL EVIDENCE

The observational analysis conducted in the previous section shows evidence of substantial and persistent levels of inattention to non-left-digits. A natural question for a profit-maximizing firm is whether these insights can lead to better pricing strategies to improve profits. We designed a natural field experiment (see Harrison and List, 2004) that was rolled out to all passengers on the Lyft platform from November 18, 2019 to January 22, 2020. To distinguish this experiment from the ongoing pricing experiments Lyft runs, we denote this field experiment as the Left-Digit Bias (LDB) experiment. Our field experiment had a simple 2×3 design and included randomization for both Standard and Shared prices. During the experimental period, Lyft ran its usual pricing algorithm to produce Standard and Shared prices. However, these prices were adjusted depending on the experimental condition, which included six cells:

1. No change Standard, No change Shared (50%)
2. Standard Down, Shared Down (10%)
3. Standard Down, No change Shared (10%)
4. Standard Down, Shared Up (10%)
5. No change Standard, Shared Down (10%)
6. No change Standard, Shared Up (10%)

13. For the purposes of this exercise, we switch to reporting elasticities instead of slopes to make comparison with prior literature easier.

No change Standard and no change Shared are conditions where the Standard and Shared prices continued to be produced by the usual Lyft pricing algorithm. The Standard Down and Shared Down conditions indicate that these prices were lowered to the nearest 99-cent value if the usual pricing algorithm produced a price that was at or just over a dollar value. Specifically, Standard Down meant that if the pricing algorithm produced a Standard price from $\$X.00$ to $\$X.09$, it was lowered to $\$(X-1).99$.[14] Finally, two conditions contain Shared Up. In these conditions, the Shared price was increased to the dollar value if it was just under the dollar value. Specifically, Shared Up meant that if the pricing algorithm produced a Shared price from $\$X.90$ to $\$X.99$, it was raised to $\$(X+1).00$.[15]

The randomization occurred at the user level. Thus, users who were assigned to a given treatment, continued to receive prices produced by that treatment throughout the experimental period. 50% of users were placed in the control condition (no change Standard and no change Shared) and the other five conditions were each assigned 10% of users.

As we discuss further below, the experimental conditions chosen are not pricing strategies that are likely to be optimal given empirical estimates we report from the observational analysis. Rather, the conditions are a very conservative move toward pricing that starts to account for left-digit bias of passengers. In this way, they represent scientifically important departures in that the field experiment allows us to validate causally the results found in the observational data while speaking to additional key aspects of left-digit bias, such as what are the effects of 99-cent pricing on long-run behavior and platform use of customers?

We focused on two primary outcome measures: profit per user (referred to internally as "net revenue" since Lyft still has to pay for other variable expenses and costs) and total fares paid per user. In addition to these two main metrics, we also compute total sessions per user and conversion rates. Decomposing primary effects into secondary effects is standard practice within Lyft and is used as a method to understand the mechanism of experiments and separate out the short- and long-term impacts of various policies.

In total, the experiment included $\sim$21 million unique passengers ($\sim$10.5M in the control group and $\sim$2.1M in each of the other five other conditions) and 177 million total sessions. In Table 4, we examine pre-experiment metrics for users in each of the experimental conditions to test for balance. An F-test for each metric is computed to test for imbalance across conditions. The statistically insignificant F-statistics provide reassurance that users were properly randomized. We also report the observation count for each condition and confirm that the randomization to each condition contained users in the expected proportions.

5.1. *Forecasts*

Before discussing our field experimental results, it is natural to ask whether the findings in our observational analysis allow us to accurately predict the effects of our various treatments. Developing forecasts for the experiment can be a useful exercise to gauge the consistency of the observational analysis and the experimental evidence. Our forecasting strategy closely follows the model described in Section 3.2. We focus on the impact of the experimental policies on the firm's business outcomes. Specifically, we forecast the impact of our treatments

14. A majority of prices were eligible for price changes, with the sole exception that if a city had a multiplicative tax, the price was not lowered. This left us with roughly 70% of sessions being eligible to be lowered. Given that 30% of sessions with an intent to treat were not actually treated, the results we produce are biased toward zero. We correct for this bias in the structural estimation that follows.

15. Given the higher profitability of the Standard product, Lyft chose to include this condition to see how left-digit bias might lead to profitable substitutions from one product to another.

TABLE 4
Balance check of experiment

	Control	Std. ↓ Sha. ↓	Std. ↓ Sha. −	Std. ↓ Sha. ↑	Std. − Sha. ↓	Std. − Sha. ↑	F-Stat (p-val)
Profit	11.71	11.68	11.70	11.71	11.72	11.71	0.51
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.77)
Fares paid	30.09	30.01	30.10	30.10	30.12	30.08	0.58
	(0.02)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.72)
Sessions	3.30	3.30	3.31	3.31	3.31	3.30	1.46
	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.20)
Std. rides	1.78	1.77	1.78	1.77	1.78	1.78	4.47
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.55)
Sha. rides	0.34	0.35	0.34	0.35	0.35	0.35	2.1
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.06)
N	10,491,428	2,097,470	2,099,645	2,096,242	2,097,968	2,097,995	

Notes: This table performs a balance check on metrics of interest from the experiment. Each row, except for the last one, corresponds to a given metric while each column, except for the last one, corresponds to a given experimental condition. The unit of analysis is the user. With each of these rows and columns, we display the mean of that metric per user in the 28 days immediately before the experiment started. Standard errors are displayed in the parentheses. The last row displays the number of users within each experimental condition while the last column displays the F-statistic (with p-value in parentheses) corresponding to the relevant metric of interest.

on total fares paid by the consumer and profits. For the profitability numbers, we make reference to the quantities C^{std}, C^{sha} which, respectively, denote the marginal cost of serving a given Standard or Shared ride. Consider fixing the prices of Standard and Shared, respectively, at $P^{std} = p^{std}$, $P^{sha} = p^{sha}$. Fix $k = std, sha$ and consider the effect of moving p^k down to $\$p^k - 0.01$. In this case, the *actual* price change is $-(¢p^k + 0.01)$, which we will denote Δ^k. Alternatively, under our model of inattention, the *perceived* price change will be $\hat{\Delta}^k \equiv -[(1-\theta)¢p^k + 0.01 + \theta 0.99]$. Similarly, consider the move of p^k up to $\$p^k + 1$. In this case, the *actual* price change, Δ^k, is $1 - ¢p^k$, while the *perceived* price change is: $\hat{\Delta}^k \equiv 1 - (1 - ¢p^k)\theta$.

For the purposes of this section, and based on the observational evidence, we take $\theta = 0.5$. Given these definitions, our forecast of fares paid after the price change can be written as

$$\begin{aligned}\mathbb{E}[(D^{std}(p^{std}, p^{sha}) + \hat{\Delta}^{std}\beta^{std,std} + \hat{\Delta}^{sha}\beta^{std,sha}) \cdot (p^{std} + \Delta^{std})\\ + (D^{sha}(p^{std}, p^{sha}) + \hat{\Delta}^{std}\beta^{sha,std} + \hat{\Delta}^{sha}\beta^{sha,sha})\\ \cdot (p^{sha} + \Delta^{sha}) \mid P^{std} = p^{std}, P^{sha} = p^{sha}]\end{aligned} \tag{8}$$

while the forecast for profits is given by

$$\begin{aligned}\mathbb{E}[(D^{std}(p^{std}, p^{sha}) + \hat{\Delta}^{std}\beta^{std,std} + \hat{\Delta}^{sha}\beta^{std,sha}) \cdot (p^{std} + \Delta^{std} - C^{std})\\ + (D^{sha}(p^{std}, p^{sha}) + \hat{\Delta}^{std}\beta^{sha,std} + \hat{\Delta}^{sha}\beta^{sha,sha})\\ \cdot (p^{sha} + \Delta^{sha} - C^{sha}) \mid P^{std} = p^{std}, P^{sha} = p^{sha}]\end{aligned} \tag{9}$$

We construct our forecasts by connecting (8) and (9) to our data. Since Δ^k, $\hat{\Delta}^k$ only depend on the cents amount in the price, in describing our forecasts, for $c \geq 0$, let Δ^k_c, $\hat{\Delta}^k_c$ be the above quantities as a function of the cent amount in the price. Additionally, to simplify notation, for $c \geq 0$, let $\Delta^k_{-c} = \Delta^k_{1-c}$ and $\hat{\Delta}^k_{-c} = \hat{\Delta}^k_{1-c}$. Our forecasts are therefore constructed using the following procedure:

TABLE 5
Forecasts of experiment results

	Fares paid	Profit
Control	+0.00%	+0.00%
Std. ↓, Sha. ↓	+0.21%	+0.16%
Std. ↓, Sha. −	+0.18%	+0.15%
Std. ↓, Sha. ↑	+0.15%	+0.14%
Std −, Sha. ↓	+0.01%	+0.03%
Std −, Sha. ↑	−0.03%	−0.01%

Notes: This table shows our forecasts of how much profits and fares paid change from different experimental conditions. Each row corresponds to a given treatment. The first column forecasts profits while the second column forecasts fares paid. The forecasts are expressed as percent changes over the status quo (*i.e.* the control condition).

1. Group prices according to the prices of Standard and Shared rounded to the nearest integer. Denote the groups by $g = (\$^{std}, \$^{sha})$.
2. Within each group, g, compute $\bar{D}_g^{std}$, $\bar{D}_g^{sha}$, respectively, as the average values of demand for Standard and Shared, let $\bar{\beta}_g^{jk}$ be the slopes of the demand curve as estimated in Section 4, and let $\bar{C}^{std}$, $\bar{C}^{sha}$ be the average cost.
3. Compute $\pi_g^{c,c'}$ as the proportion of the sessions within each group with Standard cents amounts ending in c and Shared cents amounts ending in c'. For the policy under consideration, define $\hat{\Delta}^c$, Δ^c according to the rules of the policy.
4. Forecast the fares paid after the price change within group g as

$$FP_g^{new} = \sum_{c=-50}^{49} \sum_{c'=-50}^{49} \pi_g^{c,c'} [(\bar{D}^{std} + \hat{\Delta}_c^{std} \bar{\beta}^{std,std} + \hat{\Delta}_{c'}^{sha} \beta^{std,sha}) \cdot (\$^{std} + c + \Delta_c^{std}) \\ + (\bar{D}^{sha} + \hat{\Delta}_c^{std} \bar{\beta}^{sha,std} + \hat{\Delta}_{c'}^{sha} \beta^{sha,sha}) \cdot (\$^{sha} + c' + \Delta_{c'}^{std})]$$

5. Forecast profits after the price change as

$$PR_g^{new} = \sum_{c=-50}^{49} \sum_{c'=-50}^{49} \pi_g^{cc'} [(\bar{D}^{std} + \hat{\Delta}_c^{std} \bar{\beta}^{std,std} + \hat{\Delta}_{c'}^{sha} \beta^{std,sha}) \cdot (\$^{std} + c + \Delta_c^{std} - \bar{C}^{std}) \\ + (\bar{D}^{sha} + \hat{\Delta}_c^{std} \bar{\beta}^{sha,std} + \hat{\Delta}_{c'}^{sha} \beta^{sha,sha}) \\ \cdot (\$^{sha} + c' + \Delta_{c'}^{std} - \bar{C}^{sha})]$$

6. Compute FP_g^{old}, PR_g^{old} with the above formulas, but with $\hat{\Delta}_c^k$, $\hat{\Delta}_c^k$ set to 0.
7. Forecast the percent change in fares paid as $\frac{\sum_g N_g FP_g^{new}}{\sum_g N_g FP_g^{old}} - 1$ and forecast the percent change in profits by $\frac{\sum_g N_g PR_g^{new}}{\sum_g N_g PR_g^{old}} - 1$ where N_g is the number of sessions in group g.

Empirical results of our forecasts are presented in Table 5. Relative to the control group, the largest increase in profits and fares paid is expected in the experimental conditions where the Standard product price is lowered. This is because the profit margin on the Standard product is higher than that of the Shared product. We also predict gains from lowering the Shared price and losses from raising the Shared price. Even though the profit margins are lower for the Shared rides, moving the price up to the dollar threshold causes a reduction in overall conversion that dominates the positive value to Lyft of passengers substituting to the Standard product.

As a final note, the forecasts made here are short-run in the sense that we do not take into account the potential longer-run effects of 99-cent pricing. For example, it is possible that lowering prices to just under a certain dollar value could lead customers to believe that prices, in general, at Lyft are cheaper and therefore be more likely to use the app again in the near future. One could also imagine an opposite effect on usage if 99-cent prices irritate customers or cause them to believe that Lyft is exploiting them with "gimmicky" prices. In the results below, we test for this directly by examining the average number of sessions per user by experimental condition.

5.2. *Field experimental results*

Our discussion of the field experimental results begins by visually demonstrating the impact of the experiment on Lyft profit per user. In Panel A of Figure 8, we plot the percentage difference in profit for each of the five treatment conditions relative to the control condition. We show comparisons between treatment and control in three distinct time periods. The periods were chosen to be adjacent in time and with each spanning the exact length of the experiment. Thus, the first five bars compare treatments to control during the ∼2 months prior to the experiment, the next five bars compare the treatment to control during the ∼2 month-long experiment, and the last five bars compare the treatment to control during the ∼2 months after the experiment ended.

The first five bars indicate that profit margins for the users assigned to the treatment conditions were similar to those in the control condition in the pre-period (although the Standard Down and Shared Down condition was a bit low relative to control). The next five bars illustrate that during the experiment, two of the experimental conditions (Standard Down, Shared No Change and Standard Down, Shared Up) produced profits that were significantly higher than control using conventional significance thresholds. In the final five bars, we see that in the post-experiment period, there is no statistically significant difference between the treatment conditions and the control group.

Despite the large number of users in the experiment, the point estimates for each treatment have relatively large standard errors. One of the reasons for this noise is that for the vast majority of sessions, being in one of the treatment conditions was no different than being in the control condition. For example, the treatment that lowers the Standard price only affects the price on 7% of the sessions (only 10% of sessions have a price that is within 10 cents over a dollar threshold and only 70% of sessions do not have a multiplicative tax and can be adjusted). To increase statistical power, we focus on sessions where the treatment conditions could have differed from the control condition. More concretely, we add up the profit for just sessions where at least one price was within 10 cents of the nearest dollar and in locations without a multiplicative tax.

Panel B of Figure 8 shows the profit per user for these rides relative to control in an analogous fashion to Panel A of the same figure. The percent impacts are considerably larger as expected, and we find that profit per user is significantly higher for the three treatments that include rounding Standard Down at conventional levels. We move from a visual representation to a table in order to quantify the exact effect sizes for each experimental condition.

Table 6 shows the results of each treatment relative to control for profit per user, fares paid per user, sessions, and conversion (for both Standard and Shared).[16] As a useful check, let us first focus on the impact of the experiment on Standard and Shared conversion rates. Lowering the Standard price (the first 3 rows in Table 6) increases conversion rates for the Standard product,

16. For the results presented in this table, we follow a standard variance-reduction procedure based on user-specific pre-treatment levels. See Appendix D for a discussion of our method.

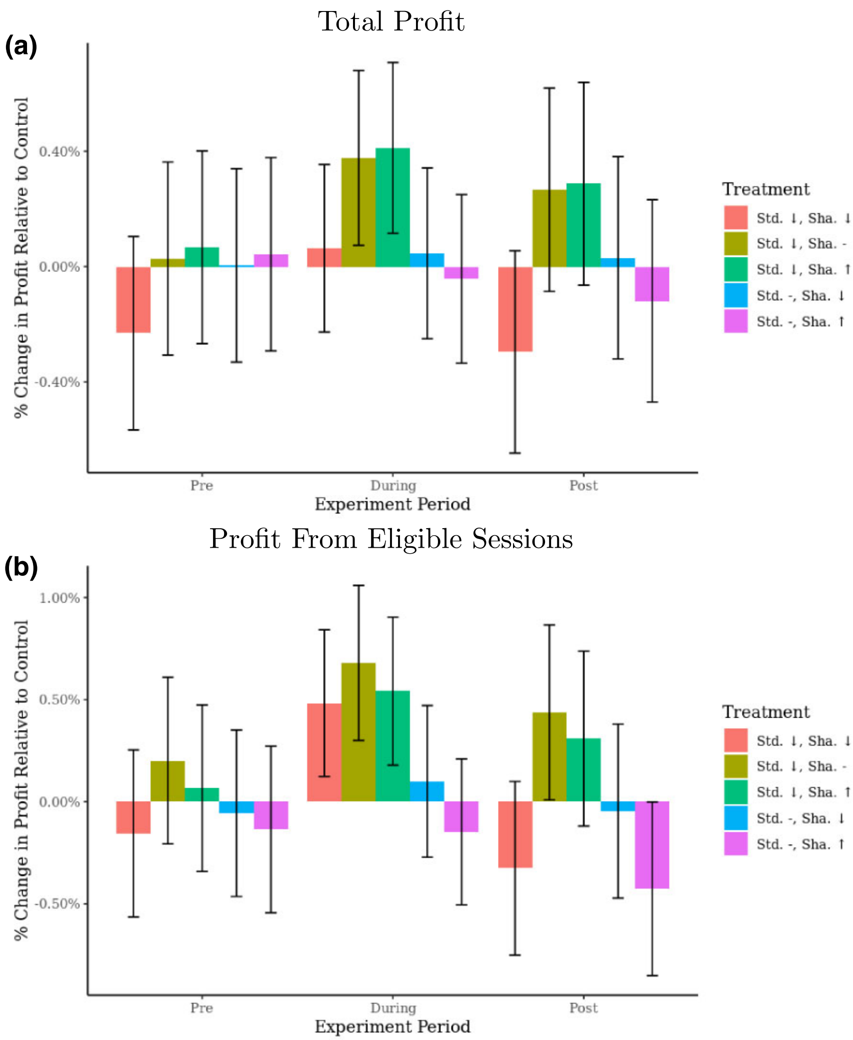


FIGURE 8

Profit in experiment

Notes: This figure plots the impact of the experimental conditions on Lyft's profit/user. Each bar shows the percent difference in the average profit per user in a given treatment group against the average profit per user in the control group, with error bars corresponding to the 95% confidence intervals. The unit of analysis is the user. Standard errors are computed by first estimating standard errors for the level of profit/user in each experimental condition and using the delta method to get standard errors for the percent change. We plot the differences for three periods: pre experiment, during the experiment, and post experiment. "Pre experiment" is defined as the time occurring just before the first treatment was assigned with time length equal to the length of the experiment. Similarly, "Post experiment" is defined as the time occurring just after the experiment was turned off with time length equal to the length of the experiment. Within each period, the bars are, in order from left to right, Standard down, Shared down; Standard down, Shared no change; Standard Down, Shared up; Standard no change, Shared down; Standard no change, Shared up. Panel (a) plots profits coming from all sessions. Panel (b) plots profit coming only from sessions where the price could have been altered by at least one of the experimental conditions.

especially when the Shared price was not also lowered or was raised. Lowering the Shared price without changing the Standard price resulted in a higher Shared conversion rate. Similarly, raising the Shared price resulted in a significantly lower Shared conversion rate.

Examining the number of sessions per user we find—somewhat surprisingly—large and significant results. All of the conditions where the standard price is lowered resulted in a statistically

TABLE 6
Treatment effects of experiment

	Std. conv.	Sha. conv.	Sessions	Fares paid	Profit
Std. ↓, Sha. ↓	0.04%	0.30%	0.21%*	0.25%*	0.20%
	(0.07)	(0.25)	(0.09)	(0.11)	(0.11)
Std. ↓, Sha. −	0.20%**	−0.12%	0.26%**	0.27%*	0.33%**
	(0.06)	(0.25)	(0.09)	(0.11)	(0.11)
Std. ↓, Sha. ↑	0.29%**	−1.00%**	0.28%**	0.46%**	0.39%**
	(0.07)	(0.26)	(0.09)	(0.11)	(0.11)
Std. −, Sha. ↓	−0.18%*	0.73%**	0.15%	0.02%	0.04%
	(0.07)	(0.25)	(0.09)	(0.11)	(0.11)
Std. −, Sha. ↑	0.07%	−1.20%**	0.01%	−0.11%	−0.05%
	(0.07)	(0.25)	(0.09)	(0.11)	(0.11)

Notes: This table shows the impact of each experimental condition on the metrics of interest relative to control. Rows correspond to the different treatments while columns correspond to various metrics. Standard/Shared conversions are defined as average Standard/Shared rides per user divided by average sessions per user. Sessions, fares paid, and profits are computed per user. The unit of analysis is the user. Each entry in the table gives the percent change in the metric over control for a given experimental condition (so if in control, Standard conversion is 0.5 and the treatment effect for a given variant is +0.20%, then Standard conversion in that variant is 0.501). Standard errors are given in parentheses. Levels in control and treatment effects are estimated using standard variance reduction techniques by using pre-period outcomes as controls in a fully interacted ordinary least squares (OLS) as described in Negi and Wooldridge (2019). For conversion metrics, standard errors are computed by taking 200 bootstrap resamples (where each *user* is resampled with replacement) and recomputing conversion on these resampled datasets. For all other metrics, standard errors are computed by using the delta method on the (heteroskedasticity and misspecification robust) standard errors from our OLS estimates and using the delta method. Additional details for how we computed standard errors can be found in Appendix D. *indicates statistical significance at the 5% level. **indicates statistical significance at the 1% level.

significant increase in the number of sessions per user (a 0.2%–0.3% increase in a number of sessions). This is consistent with the theory that passengers who are in a condition where the Standard price is lowered remember their experience as better or the prices being lower, which leads them to increase their usage of Lyft in the future.

The increases in conversion and sessions lead directly to increases in fares paid and profit. The largest increase in profit (0.39%) comes from the condition where Standard prices are lowered and Shared prices are increased. But, we see a similar profit increase in the condition where only Standard prices are lowered.

We now compare our results to our forecasts. In Table 7, we report the ratio of treatment effects actually observed in the experiment to the forecasts along with standard errors. The effects on profit that we find are largely directionally consistent with our forecasts. More quantitatively, we find that while our estimates are quite noisy, in general, it appears that our forecasts tended to underestimate the effect of our rounding treatments. This is consistent with the fact that the treatments that lowered prices appear to have longer-run effects on Lyft demand by increasing the number of sessions taken by treated passengers, which were not accounted for in the original forecasts.

5.3. *Structural estimation*

Using experimental variation, we can once again structurally estimate the inattention parameter as described in (3). The approach we take continues to build on the empirical model introduced in Section 3.2. Thus, similar to the section using observational data, our basic estimation approach heuristically boils down to comparing the size of a "discontinuity" with the slope from larger

TABLE 7
Forecasts compared to experiment results

	Fares paid	Profit
Std. ↓, Sha. ↓	1.19	1.25
	(0.44)	(0.55)
Std. ↓, Sha. −	1.22	2.2
	(0.61)	(0.73)
Std. ↓, Sha. ↑	3.07	2.79
	(0.56)	(0.79)
Std. −, Sha. ↓	2.00	1.33
	(11.00)	(3.67)
Std. −, Sha. ↑	3.67	5.00
	(3.67)	(11.00)

Notes: This table compares how the experiment results compare to our forecasts. Each row corresponds to a given treatment. The first column divides the actual experiment effect on Fares Paid by the corresponding forecast. The second column divides the actual experiment effect on Profits by the corresponding forecast. Ratios close to 1 reflect accurate forecasts. Standard errors are given in parentheses and are constructed under the assumption that the forecasts are non-random.

price changes. Unlike the observational section, however, because our dataset is a more manageable size, we are able to perform our estimation in a manner that allows us to test formally the quantitative predictions of the left-digit bias model.

Details of the estimator are presented in Appendix E. At a high level, the estimator works by translating the implications of the left-digit bias model into a set of moment conditions, which we can estimate using a generalized method of moments (GMM) approach. In our baseline model, we assume that consumers face a single inattention parameter governing both attention to the cents in Standard price and attention to cents in Shared price and that this inattention is uncorrelated with the slope of the demand curve. The restrictions placed by Corollary 3 imply that these moment conditions are over-identified, which allows us to partially test our inattention model.

In Table 8, we report the parameters estimated using the GMM outlined above. As shown in column 1 and consistent with the observational evidence, we find an average inattention parameter of 0.45. Along with inattention, we also produce estimates for the derivatives (two own-price derivatives and two cross-price derivatives) of the latent demand curve. The basic model of inattention we have been using does not take a view on how the latent demand curve should be shaped. However, we do obtain meaningful restrictions on the shape of the latent demand curves under the additional assumptions that modulo being inattentive to prices, consumers are utility maximizers and that there is no income effect. In particular, these assumptions imply that the matrix of price derivatives of latent demand should be symmetric and negative semi-definite, which we do observe in our estimates of the price derivatives. Under our additional modeling assumption that latent demand curves are linear, this implies that cross-price slopes should be equal to one another, and the matrix of own and cross-price slopes is negative semi-definite. Our results on the slopes of the latent demand curve are unable to reject these predictions.

In column 2, we re-run the estimation performed in column 1, but allow for inattention to differ between Standard and Shared prices. Estimating the parameters separately, we find inattention of Standard prices to be slightly higher, 0.49, but still similar to the inattention estimate of Shared prices, 0.42.

Lastly, we explore heterogeneity in the inattention parameter. Using the observational data, we did not find evidence that inattention varied much by passenger or ride type. One potential exception was that passengers in sessions that had both the Standard and Shared products

TABLE 8
Inattention parameter from experimental data

	Std. and Sha.	Std. and Sha.	Std. only
Std. to Std.	−0.036	−0.0358	−0.031
	(0.0009)	(0.0009)	(0.001)
Std. to Sha.	0.012	0.012	
	(0.0005)	(0.0005)	
Sha. to Std.	0.011	0.011	
	(0.001)	(0.001)	
Sha. to Sha.	−0.024	−0.025	
	(0.0007)	(0.0008)	
Inattention (θ)	0.45		0.29
	(0.02)		(0.04)
Inattention to Std. (θ_{std})		0.49	
		(0.03)	
Inattention to Sha. (θ_{sha})		0.42	
		(0.03)	

Notes: This table shows estimates for inattention along with estimates for the derivatives of Standard and Shared demand to changes in Standard and Shared prices. The sample used for estimating inattention here is constructed by taking the first session where a user *could have* had their price altered by any of the treatment variants (so that each individual in the experiment had at most one observation, and no long-run effects of treatment are present within the sample). A derivative labeled X to Y means the derivative of Y demand with respect to exogenous changes in X price. These parameters are estimated by constructing a system of observable moments as described in Appendix E and using a two-step GMM estimator in the case where the system of moments is over-identified (first two columns) and a methods of moments estimator when the moments are just identified (last column). Standard errors are reported in parentheses. The first column produces estimates for inattention under the assumption that Standard and Shared inattention are identical. The second allows inattention to Standard price to differ from inattention to Shared price. The third column estimates inattention in sessions that only displayed a Standard price.

appeared to be more inattentive compared to those in sessions with just the Standard product. In our experimental sample, we have the opportunity to test if this effect is robust. By taking the subset of the moments that only feature the Standard option, we are able to identify inattention. Empirical results of this analysis are shown in column 3 of Table 8. We find an inattention parameter of 0.29, which is similar to what we found in the observational sample, and is once again statistically different from the inattention parameter for sessions with Standard and Shared at conventional levels.

5.4. *Overall value to Lyft*

Using the field experimental results, we estimate the overall value to Lyft from the most effective experimental treatment to be a 0.39% increase in profit per user. The treatments from our experiment, however, were very cautious movements toward adjusting prices in response to passengers with left-digit bias. What would be the impact on total profit if Lyft chose to pursue a more aggressive 99-cent pricing strategy? Answering this question fully would require us to model carefully Lyft's objective function, but just our estimate of θ allows us to provide a rough sense of the magnitude of the potential gains.

The key idea is that inattention implies a disconnect between actual and perceived prices. A firm can use this disconnect to move prices such that on average, perceived prices remain constant while actual prices increase. Assuming an inattention parameter of θ, and assuming that the cents amount in prices is uniformly distributed (which, as mentioned earlier, appears to be a good approximation), this can be accomplished by moving all prices with cents amounts below a threshold t down to the 99-cent ending price below while moving all prices with cents amount above t up to the 99-cent ending price above. The threshold such that this keeps average perceived prices fixed is approximately $t=\frac{1-\theta}{2}$. Because perceived prices are fixed on average, total demand, and hence total costs, stay constant. Letting Q be the quantity sold and letting β be the (magnitude of the) slope of the latent demand curve, the change in average price for goods sold is given by

$$\underbrace{\frac{1}{Q}\int_t^1 [Q-(1-\tau)(1-\theta)\beta](1-\tau)\,\mathrm{d}\tau}_{\text{Contribution from price increases}} - \underbrace{\frac{1}{Q}\int_0^t [Q+\tau(1-\theta)\beta]\tau\,\mathrm{d}\tau}_{\text{Contribution from price decreases}}$$

This formula can be rearranged to

$$\frac{1}{2}-t+\left[\frac{t^3-(1-t)^3}{3}\right](1-\theta)\frac{\beta}{Q}$$

or plugging in $t=\frac{1-\theta}{2}$,

$$\frac{\theta}{2}+\left[\frac{(1-\theta)^3-(1+\theta)^4}{24}\right](1-\theta)\frac{\beta}{Q}$$

The first term of the above expression represents the change in average price from our proposed heuristic policy if there was no accompanying behavioral change. The second term corresponds to the change in average price due to the fact that higher prices lead to less demand while lower price leads to more demand, thus shifting the price change in favor of the lower prices. This second term implies that our heuristic strategy may even decrease profits for values of θ sufficiently small, but we find that plugging in $\theta=0.5$, this term is negligible for any reasonable parameterizations of Lyft's demand curve.[17] Assuming that the second term is negligible, one can therefore approximate the increase in profits from such a heuristic policy by $\frac{Q\theta}{2}$ where Q is quantity sold. Given an estimate of θ, and some range of plausible values for β, the above discussion provides a back-of-the-envelope approach to estimate the potential gains of 99-cent pricing not only for Lyft, but for *any* company where quantity data are available. In the case of Lyft, our ballpark estimates amount to an estimated potential gain of $\frac{\theta}{2}=\$0.25$ per ride from adopting 99-cent pricing. Given the roughly 650M rides on Lyft's platform in 2019, this amounts to an increase in profits of around \$160M.

6. DISCUSSION

6.1. *Why was not Lyft optimizing for left-digit bias?*

Given the magnitude of our results, a reader may wonder why Lyft was not optimizing for left-digit bias in their pricing policy prior to our experimentation. Some institutional details might begin to shed light on this question.

Our research was, in fact, not the first experiment to explore the idea of 99-cent pricing with rideshare. Specifically, an earlier experiment had attempted to set all prices in a small handful of cities at 99 cents but was unable to detect a significant effect on ridership and profitability. This

17. Specifically, suppose $\beta/Q=1$. Even in this case, the second term amounts to roughly -7¢. Moreover, note that $\beta/Q=\varepsilon/P$ where ε is the elasticity of demand and P is the price level. Thus, $\beta/Q=1$ would imply an elasticity of around 17, an order of magnitude larger than any elasticity estimated in rideshare to date.

experiment was motivated by an observation on Lyft's part that 99-cent pricing was common practice in the industry without necessarily having a justification for *why* such a policy might be a good idea. However, given our estimated inattention parameter of $\theta \approx 0.5$, a power calculation we performed for Lyft in order to convince them to run our experiment showed that this previous experiment simply did not have a sufficiently large sample size to detect the desired effects. Without a theoretical framework to guide this analysis, it was difficult for Lyft to recognize that the failure to detect an effect was due to insufficient data, not necessarily due to the absence of an effect.

Nonetheless, in light of the stark observational trends shown in Figures 3 and 4, one might wonder why Lyft's suite of sophisticated machine-learning tools was unable to pick up such a sizeable discontinuity in demand. Here again, the institutional details of how the machine learning system was set up at Lyft stacked the cards against finding an effect. Specifically, in order to deal with the inherently high-dimensional problem of predicting demand based on a large number of covariates, Lyft's machine learning system must impose a number of both explicit and implicit prior beliefs on the regularity of the relationship between those covariates and demand.[18] A key restriction that is especially important for our purposes is that Lyft's machine learning systems assumed from the outset that demand curves were locally well approximated by a constant elasticity functional form conditional on covariates. This sort of regularity restriction was helpful for Lyft because the imposed smoothness of the demand curve helped stabilize the final price that the pricing algorithm produced, but it clearly biased the overall system against finding evidence of left-digit bias. Our view in light of this discussion is that these modeling choices by Lyft seemed ex ante to be a sensible approximation for helping the machine learning algorithm converge. Only ex post (*i.e.* given what we know from this present study) was it obvious that these modeling choices prevented Lyft from finding what turned out to be an important effect.

This discussion highlights an important point for economists working in data rich settings more generally. Even in the presence of highly sophisticated algorithms, seemingly small details about how the system is designed can have important effects on what is and is not easy to detect. Economic theory can thus play a key role in ensuring that these algorithmic systems work well, even as machine learning tools begin to automate some of the aspects of the data analysis workflow.

6.2. *Relationship of our estimates to prior literature*

In this paper, we estimate an inattention parameter in two different settings. While the results are fairly similar across our two settings, the inattention parameter we estimate is much larger than estimates in the previous literature. Specifically, both Lacetera *et al.* (2012) (henceforth LPS) and Strulov-Shlain (2021) (henceforth S) find inattention parameters ranging between $\theta = 0.2$ and $\theta = 0.3$, roughly half the size of the estimates we produce.

In this section, we use the lens of the model developed in Section 3.2 to discuss formally the unique advantages of our setting over previous studies in quantifying a precise inattention parameter. Specifically, Corollary 2 highlights three main data requirements to back out an inattention parameter. First, we need a credible estimate of the causal effect of what happens when the price crosses a dollar threshold. Second, we need a credible estimate of the causal effect of price changes spanning multiple dollar amounts. Third, we must divide these two causal effects, requiring that the estimates from the first two steps correspond to comparable populations.

18. Note that this issue is not unique to Lyft. Many standard machine learning textbooks (*e.g.* Györfi *et al.*, 2002; Murphy, 2012) discuss a set of results known as "no free lunch" theorems, which formalize the notion that learning a relationship can be arbitrarily difficult without imposing some prior restrictions on the data-generating process.

With respect to the first requirement, both LPS and S as well as our own "observational" section all use some form of regression discontinuity argument to justify the claim that the size of the discontinuity can be interpreted causally. Relative to these prior studies, an advantage of our setting is that we do not find evidence of bunching of prices below integer amounts. This allows us to avoid concerns that prices end up just below integer amounts due to deliberate action on the part of suppliers/firms. This allows us to have a high degree of confidence when estimating discontinuities. Indeed, both prior studies discuss at some length how to empirically deal with the observed bunching (*e.g.* use donut regression discontinuity approaches). Moreover, in our experimental section, by design, we side-step this assumption altogether, as we can effectively use the experimental variation as an instrument for crossing the discontinuity. The fact that the two approaches yield similar results provides reassuring evidence that our regression discontinuity indeed identifies the relevant causal effect.

The second criterion is arguably the biggest advantage of our setting. Specifically, even in our "observational" setting, the causal effect of large price changes is estimated using experimental variation. By contrast, LPS and S both must use a bevy of controls to mitigate the effect of selection bias as much as possible. LPS justify their causal interpretation by appealing to the fact that they have a rich set of controls while S deals with these concerns by estimating regressions at the chain level and including a number of fixed effects. These approaches strike us as reasonable, and therefore our approach serves to complement their work by using a different identification strategy.

The third criterion is the hardest for our study to contend with. For example, in the LDB experiment, only prices with cent amounts within 10 cents of integers are rounded. We estimate our inattention parameter by comparing this group to the price sensitivities measured in the general population. We formally justify this assumption by making the argument (which we view as plausible) that the pricing algorithm before our experiments was relatively agnostic to the precise cent amount of the final price, and hence there is unlikely to be systematic selection on cent amount. Arguably, the identification of the inattention parameter in LPS and S relies on the same implicit assumption.

Given this discussion, we are left with a puzzle, which our data alone are unfortunately not equipped to answer. Is the difference between our inattention parameter and those estimated by LPS and S due to the cleaner empirical setting available to us or fundamental differences in the level of attention in the market we study? S argues that if anything, his estimate of the slope of the latent demand curve is biased downward in magnitude. Correcting such a bias would cause our estimates to diverge even further. Insofar as pricing endogeneity is due to unobserved components that affect both demand and costs, the direction of the bias may apply more broadly, and hence, one might imagine that LPS's inattention parameter is also, if anything, biased upward. Ultimately, such questions about how inattention varies across circumstances require more inattention parameters to be estimated in more settings. The modeling techniques provided in this paper can aid in interpreting this future work by helping to clarify the key assumptions at play in making these inattention estimates credible.

6.3. *Mechanism and long-term value to Lyft*

A limitation of our large-scale field experimental setting is that while we are able to precisely and cleanly quantify the degree of left-digit bias, we are unable to learn much about the underlying mechanism. Are we observing rational inattention as Lyft riders trade-off decision accuracy and cognitive costs? If so, how should cognitive costs be parameterized? These questions about mechanism are not only important for our scientific understanding of how left-digit bias works, but also have policy relevance for Lyft. For example, if consumers behave according to a model like that of Basu (1997) and eventually learn the distribution of cents, we might find that in the long-run, 99-cent pricing is ineffective at raising Lyft's profits.

Within our roughly ten-week experiment, we found no evidence that the effect of our treatments became more negative over time (and in fact, we find evidence that the effect became more *positive* over time). However, ten weeks may still be too short for long-run effects to arise and it is possible that our treatments did not set enough prices at 99-cent ending numbers for consumers to notice. It would be interesting to see if the effects of 99-cent pricing begin to disappear either if more time is allowed to elapse or if more 99-cent ending prices are served.

7. CONCLUSION

In this paper, we explore left-digit bias in a large and important industry. We document that, prior to our study, Lyft was not employing a 99-cent pricing strategy with their pricing algorithm. Using observational data alongside a natural field experiment, we show that demand curves contain large and important discontinuities at round dollar values. The results showcase the robustness and importance of left-digit bias when considering a pricing strategy.

In a practical sense, we establish that Lyft can increase profits substantially by pricing in a way that accounts for the left-digit bias of passengers. But what about passengers? What are the welfare implications for the consumers if Lyft adopts increasingly aggressive 99-cent pricing strategies? The answer is not clear. Consider a consumer who is choosing a bundle of goods that maximizes utility subject to a budget constraint. If Lyft begins pricing rides at 99-cent marks and the consumer is inattentive, then our results suggest that they will start consuming more Lyft rides as part of their overall consumption bundle. However, perhaps they were underconsuming Lyft rides to begin with; for example, if the prices of most or all of the other goods in their consumption bundle were being priced at 99-cent endings, then the consumer was likely using Lyft too infrequently prior to Lyft adopting 99-cent pricing. This point is made in a more formal fashion by Basu (1997). He argues that if all firms price with a 99-cent ending, then it is not irrational for consumers to be inattentive and it is firms, not consumers, that experience welfare losses relative to a world with no inattention.

Our results highlight how consumers can exhibit behavioral biases and how firms often fail to respond to the behavioral biases of their customers. The size and robustness of the effects that we find point to the importance—at least in the case of left-digit bias—of examining decisions and markets through a behavioral lens, even in cases where the market is quite competitive.

APPENDIX

A. Proofs

A.1. *Proof of Proposition 1*

Consider prices around some integer dollar amount \$. Let $¢ > 0$. Then $D(\$ + ¢; \theta) = \alpha - \beta(\$ + (1-\theta)¢)$ while $D(\$ - ¢; \theta) = \alpha - \beta(\$ - 1 + (1-\theta)(1-¢))$. Then $\lim_{¢\to 0} D(\$ + ¢; \theta) = \alpha - \beta\$$ while $\lim_{¢\to 0} D(\$ - ¢; \theta) = \alpha - \beta(\$ - \theta)$ and these limits differ.

A.2. *Proof of Proposition 2*

The result follows immediately from the expressions for $D(\$ + ¢; \theta)$ and $D(\$ - ¢; \theta)$ in the proof of Proposition 1.

A.3. *Proof of Proposition 3*

Note that $s_1 \cdot (\text{¢}_1 + \text{¢}_2) = D^*(\$ + (1-\theta)\text{¢}_1) - D^*(\$ - (1-\theta)\text{¢}_2 - \theta)$, so $\lim_{\text{¢}_1,\text{¢}_2 \to 0} = D^*(\$) - D^*(\$ - \theta)$. Similarly, $s_2 = D^*(\$ + (1-\theta)\text{¢}_3) - D^*(\$ - (1-\theta)\text{¢}_3 - \theta)$. Then by the fundamental theorem of calculus,

$$\lim_{\text{¢}_1,\text{¢}_2 \to 0} s_1/s_2(\text{¢}_1 + \text{¢}_2) = \frac{\int_{\$-\theta}^{\$} (D^*)'(t)\,\mathrm{d}t}{\int_{\$-(1-\theta)\text{¢}_3-\theta}^{\$+(1-\theta)\text{¢}_3} (D^*)'(t)\,\mathrm{d}t} = \theta \cdot \frac{\frac{\int_{\$-\theta}^{\$} (D^*)'(t)\,\mathrm{d}t}{\theta}}{\int_{\$-(1-\theta)\text{¢}_3-\theta}^{\$+(1-\theta)\text{¢}_3} (D^*)'(t)\,\mathrm{d}t}$$

A.4. *Proof of Proposition 4*

WLOG, we can focus on the demand for good 1. We compute $D_1(\$ + \text{¢}, p_2; \theta) = \alpha - \beta_{1,1}(\$ + (1-\theta_1)\text{¢}) + \beta_{1,2}(\lfloor p_2 \rfloor + (1-\theta_2)(p_2 - \lfloor p_2 \rfloor))$, $D_1(\$ - \text{¢}, p_2; \theta) = \alpha - \beta_{1,1}(\$ - 1 - (1-\theta_1)(1-\text{¢})) + \beta_{1,2}(\lfloor p_2 \rfloor + (1-\theta_2)(p_2 - \lfloor p_2 \rfloor))$, $D_1(p_1, \$ + \text{¢}; \theta) = \alpha - \beta_{1,1}(\lfloor p_1 \rfloor + (1-\theta_1)(p_1 - \lfloor p_1 \rfloor)) + \beta_{1,2}(\$ + (1-\theta)\text{¢})$, $D_1(p_1, \$ - \text{¢}; \theta) = \alpha - \beta_{1,1}(\lfloor p_1 \rfloor + (1-\theta_1)(p_1 - \lfloor p_1 \rfloor)) + \beta_{1,2}(\$ - 1 + (1-\theta)(1-\text{¢}))$.

A.5. *Proof of Proposition 5*

WLOG, we focus on the case where $j = std$ and p^{std} is being varied. By our assumptions,

$$\begin{aligned}\mathbb{E}[Y_i^{std}(\$^{std} + \text{¢}, p^{sha} | U_i = u, \theta_i = t] &= \alpha_{std}(u,t) + \beta^{std,std}(u,t)(\$ + (1 - t^{std})\text{¢}) \\ &\quad + \beta^{std,sha}(\$ p^{sha} + \text{¢} p^{sha})\end{aligned}$$

while

$$\begin{aligned}\mathbb{E}[Y_i^{std}(\$^{std} - \text{¢}, p^{sha} | U_i = u, \theta_i = t] &= \alpha_{std}(u,t) + \beta^{std,std}(u,t)(\$ - 1 + (1 - t^{std})(1 - \text{¢})) \\ &\quad + \beta^{std,sha}(\$ p^{sha} + \text{¢} p^{sha})\end{aligned}$$

The result then follows by taking limits as $\text{¢} \to 0$.

A.6. *Proof of Corollary 2*

WLOG, it suffices to consider the case where the price being varied is Standard. By the law of iterated expectations and Proposition 5,

$$\lim_{p \to (\$^{std})^+} \mathbb{E}[Y_i^j(p, p^{sha}) | U_i = u] - \lim_{p \to (\$^{std})^-} \mathbb{E}[Y_i^j(p, p^{sha}) | U_i = u] = \mathbb{E}[\beta^{j,std}(u, \theta_i)\theta^{std} | U_i = u]$$

We can rewrite the RHS as

$$\mathbb{E}\left[\frac{\beta^{j,std}(u,\theta_i)\theta_i^{std}}{\mathbb{E}[\beta^{j,std}(u,\theta_i)]} \,\middle|\, U_i = u \right] \mathbb{E}[\beta^{j,std}(u,\theta_i) \mid U_i = u] = \mathbb{E}[W_i^{j,std}\theta_i^{std} \mid U_i = u]\mathbb{E}[\beta^{j,std}(u,\theta_i) \mid U_i = u]$$

A.7. *Proof of Corollary 3*

Again, WLOG, we focus on the case where Standard price is varied. Under the assumption about the covariance, both sides of the equation reduce to $\mathbb{E}[\theta_i^{std} \mid U_i = u]$ and hence are equal.

B. Robustness checks for reduced form evidence

In Figure A2, we plot the distribution of the cent amount in Standard and Shared prices. The distribution of Standard prices looks approximately uniform, although it appears that smaller cent amounts are slightly more likely than larger cent amounts. We find no evidence of bunching near 99-cent prices. The distribution of Shared prices looks somewhat strange. Specifically, we find that prices ending in multiples of 5 are over-represented, with 10-cent ending prices being especially over-represented. Conversations with engineers suggest that this was the result of rounding that occurred in an older version of the part of the pricing algorithm that was specifically responsible for determining Shared price. By the time of our experiment, this issue had already been resolved, as a newer version of the pricing algorithm (that did not round in this way) was put into production. In Figure A1, we plot the distribution of prices up to \$100. We find that prices below \$10 do not follow as smooth a distribution as prices above \$10. In the main text, we therefore choose to remove observations from these lower prices.

In Figure A3, we show that the same basic results as in the body of the paper continue to hold in Standard only regions, where the interpretation of the observed discontinuities in conversion is easier. Specifically, in Figure A3(a), we replicate the finding in Figure 2 that the distribution of Standard prices is smooth. In Figure A3(b), we replicate the result in 3 that demand for Standard drops consistently right as prices cross an integer threshold.

C. Justification for using actual price when dealing with multiplier-based price variation

In the text, we identify inattention by comparing the size of the discontinuity at each dollar amount to the elasticity of demand resulting from exogenous variation in price from the ongoing pricing experiments. The implicit assumption is that this elasticity is equal to the elasticity of the latent demand curve, which is not directly observable. Here, we show that this assumption is reasonable. In the observational setting, we use 2SLS on log price as our means of identifying semi-elasticities of demand. As a simplification, however, we use the fact that to first order, $\log(p) \approx \log(p^*) + \frac{1}{p^*}(p - p^*)$ to justify focusing on the regression of demand on unlogged price. To bound the bias of the regression estimator, we just need to bound the bias of the first stage relationship (since this is the only part of the estimator where actual price factors in). This bias is given by

$$\mathrm{Cov}(p, m) - \mathrm{Cov}(\hat{p}(p; \theta), m) = \theta \cdot \mathrm{Cov}(p - \lfloor p \rfloor, m) \tag{10}$$

The covariance above therefore gives us the relative bias[19] of 2SLS due to regressing on *actual* instead of *perceived* price in identifying the semi-elasticities of the latent demand curve. In Figure A4, we plot the norm of $\mathrm{Cov}(p - \lfloor p \rfloor, m)$ across dollar amounts, and find that it is negligible (on the order of 10^{-4}).

D. Variance reduction procedure

For each non-conversion metric, we estimate the result of a full regression adjustment as described in Negi and Wooldridge (2019) to help reduce variance. Specifically, we estimate equations of the form

$$Y_i = \beta_0 + \beta_1 \cdot D_i + \beta_2 \cdot (Y_i^{pre} - \bar{Y}_i^{pre}) + \beta_3 \cdot D_i \cdot (Y_i^{pre} - \bar{Y}_i^{pre}) + \varepsilon_i \tag{11}$$

19. By relative bias of b, we mean that if we estimated inattention on only sessions with exactly this fixed Δp, our estimate of inattention will at most be biased $u\%$ of the true inattention.

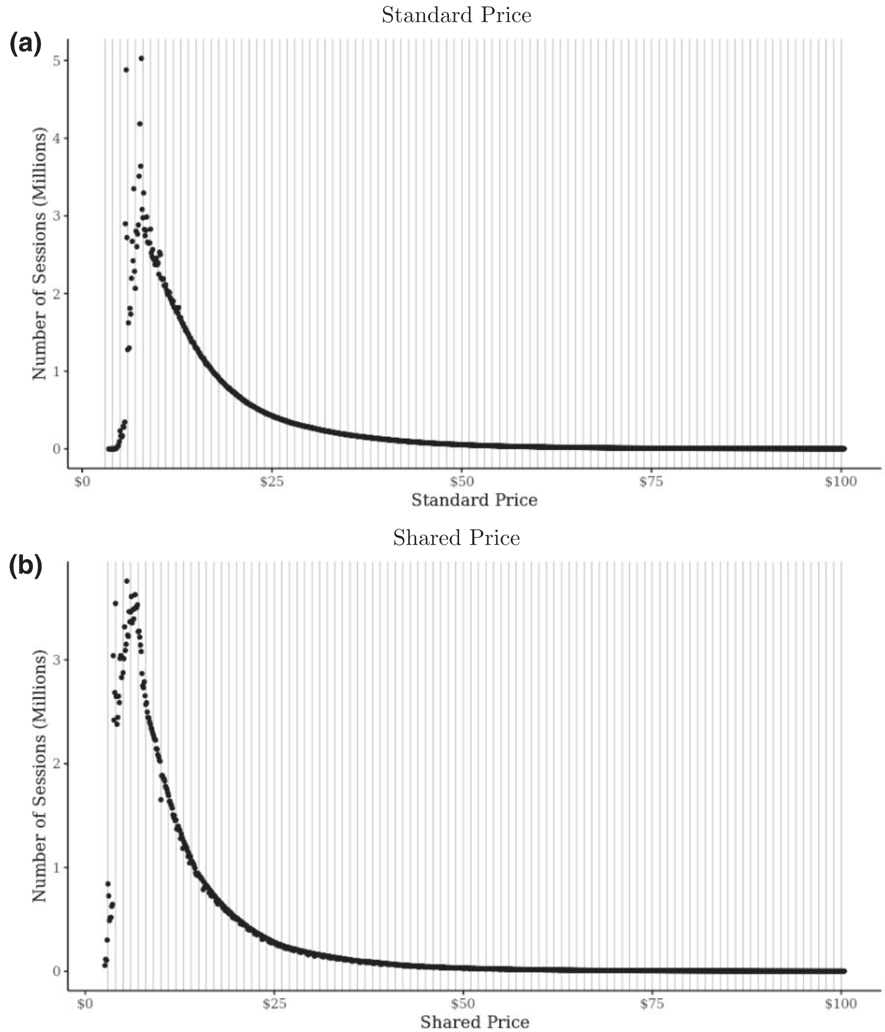


FIGURE A.1
Distribution of prices

Notes: This figure plots the distribution of Standard and Shared prices. The sample used to construct this figure is all Lyft pricing sessions with Standard and Shared prices less than \$100.05 from February 2019 through August 2019 inclusive. We additionally exclude sessions that contained prices for Lyft's Shared Saver product, because the way these sessions were priced makes the subsequent analysis difficult to interpret. Each point above corresponds to a price truncated after the first decimal (so, for example, a price of \$15.07 becomes \$15.00). Panel (a) plots the number of sessions for each truncated Standard price while Panel (b) plots the number of sessions for each truncated Shared price. The vertical lines mark integer prices.

where D_i is a dummy for treatment, and Y_i^{pre} is the value of the metric measured on an equal length period before the experiment started. Following the results of Negi and Wooldridge (2019), under this specification, $\hat{\beta}_0$ is a consistent estimate of the level of the metric in control, $\hat{\beta}_1$ is a consistent estimate of the treatment effect, and the vector $(\hat{\beta}_0, \hat{\beta}_1)$ has lower asymptotic variance than the simple difference in mean estimate. To turn the result of this estimate into a percent change, we report $\hat{\beta}_1/\hat{\beta}_0 - 1$ and compute standard errors using the delta method.

Formally, conversion is defined as #rides/#sessions. To compute conversion impacts, we therefore estimate a regression in the form of equation (11) twice. First, we estimate the regression using rides as the outcome. Second, we estimate the regression with sessions as the outcome.

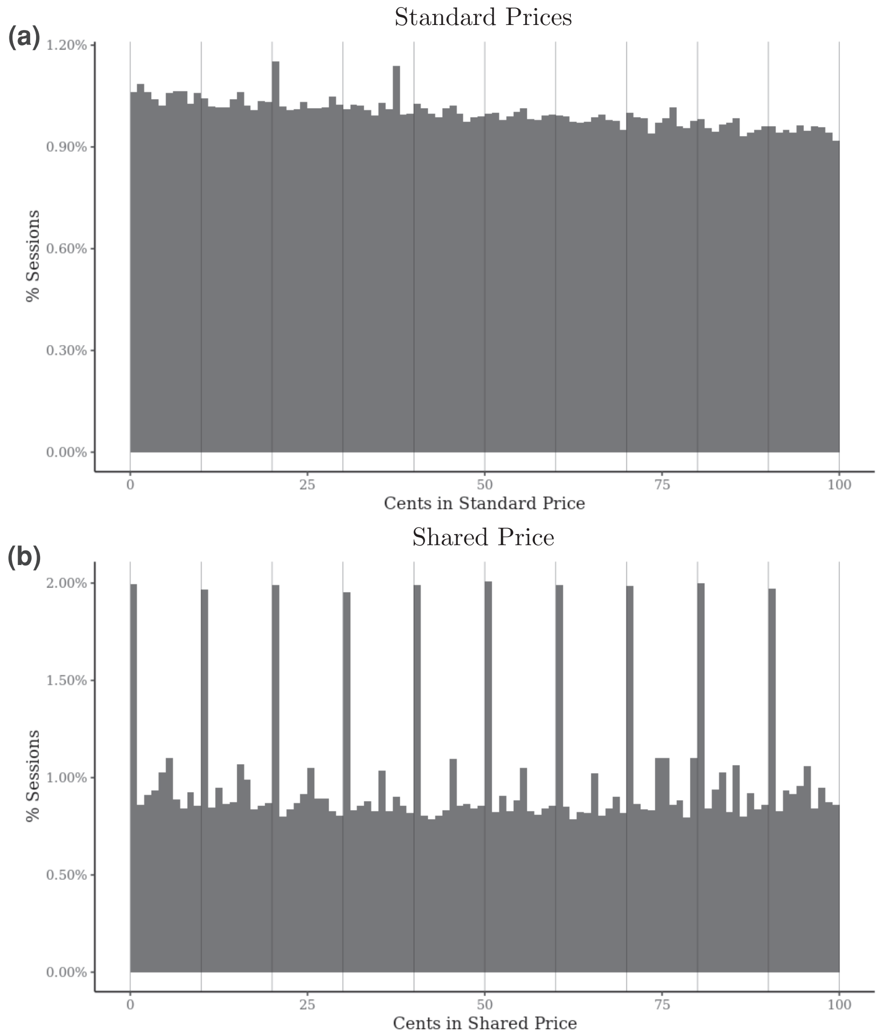


FIGURE A.2
Distribution of cents in prices

Notes: This figure plots the distribution of cent amounts in prices within our observational data. The sample used here is identical to the sample used in Figure 2 (all Lyft pricing sessions with Standard and Shared prices between \$9.50 and \$30.05 from February 2019 through August 2019 inclusive). Each bar corresponds to a single cent amount, with height representing the percent of sessions in our sample whose prices end with that cent amount.

Let the point estimates for the two regressions be, respectively, $\hat{\beta}_i^r, \beta_i^s$ for $i = 0, 1$. Plugging these into the definition of conversion, we have that our point estimate for percent impact on conversion is given by

$$\frac{(\hat{\beta}_0^r + \hat{\beta}_1^r)}{\hat{\beta}_0^s + \hat{\beta}_1^s} \cdot \frac{\hat{\beta}_0^s}{\hat{\beta}_1^r} - 1 \tag{12}$$

We compute standard errors for this regression via a bootstrap. In principle, one could get standard errors for this estimator by fitting a seemingly unrelated regression and applying the delta

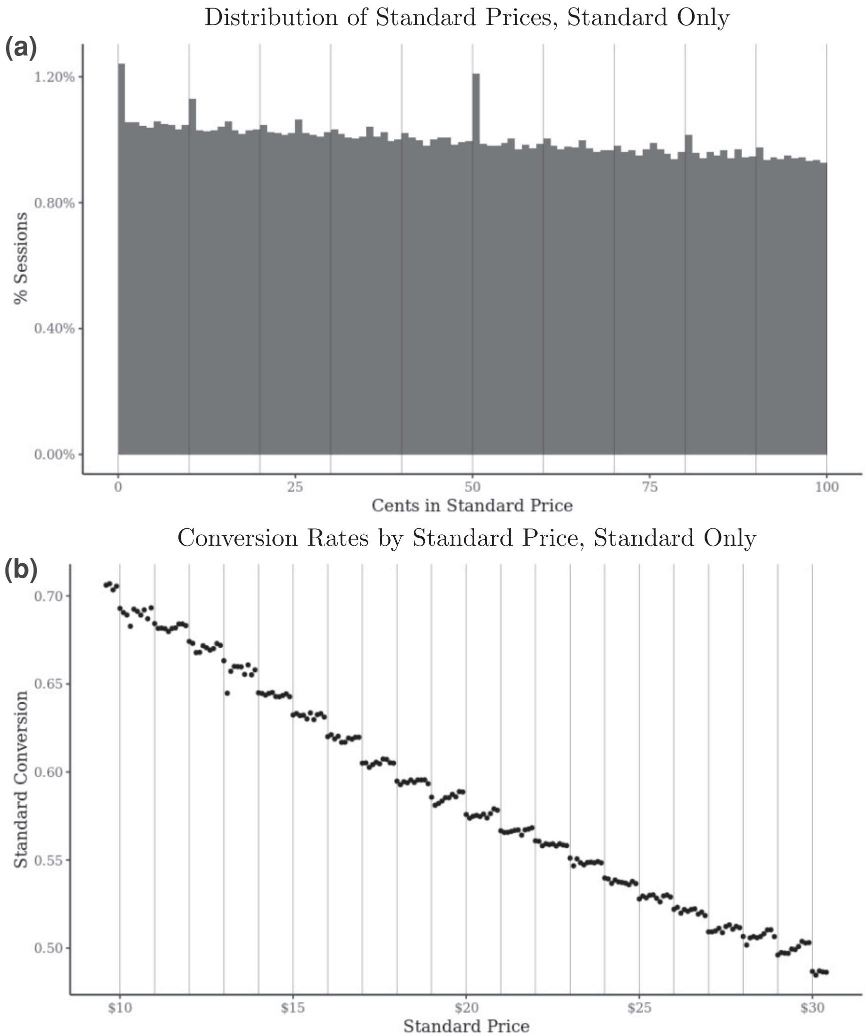


FIGURE A.3
Reduced form plots, standard only

Notes: This figure replicates the findings in Figures 2 and 3, but restricts attention to sessions where only Standard ride types were present. In Panel (b), we plot conversion to Standard for each price truncated to the first decimal. In Panel (a), we plot the distribution of Standard prices.

method. The bootstrap approach was taken instead because we could not find a software package that was computationally efficient enough able to run this on the 20 million observations we have. Specifically, we draw with replacement from our dataset of users in the experiment and implement the above for 200 draws.

E. Structural estimation details

In this appendix, we describe in more detail an estimator for inattention given the variation from the LDB experiment. Because it allows us to ignore issues associated with long-run effects of our experiment, we will first subset our data. First, the price multipliers are randomized on a

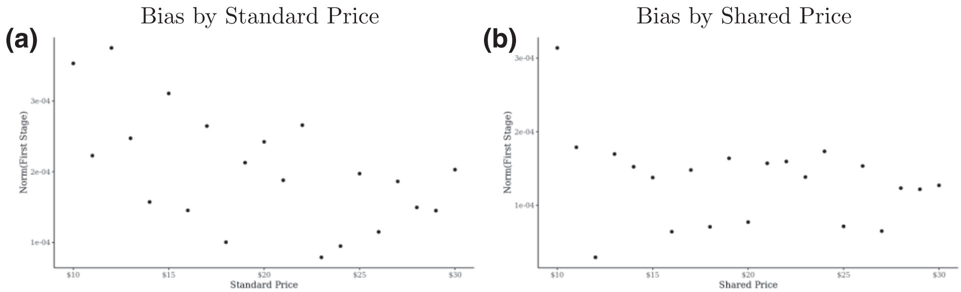


FIGURE A.4
First stage bias in 2SLS

Notes: This figure plots estimates of the first stage bias from using actual instead of perceived prices as described in equation (10) in Appendix C. Each point corresponds to an integer price, with the y-axis representing the size of the relative bias from using realized instead of perceived prices to estimate elasticity of latent demand as described in equation (10).

session-by-session basis while the experiment was randomized at the user level. We will therefore restrict our attention to the first session taken by a given user where the session could have been rounded if only the passenger happened to be in the right treatment arm.[20] We do this because this represents the first point of contact an individual has with Lyft where treatment could have possibly had a material impact on the user's experience. Second, like in the observational setting, we will focus mostly on settings where individuals had access to Standard and Shared, but not Shared Saver. As before, excluding sessions where the user had a Shared Saver price is important because there is a tight algorithmic link between the exogenous variation in Shared and Shared Saver, which makes it difficult to separate out price sensitivity parameters.

We also slightly modify Assumption 4′ to account for the additional pricing variation introduced by the LDB experiment. As before, this assumption is merely formalizing our institutional knowledge. Specifically,

Assumption 4′. Prices are determined by a formula

$$\begin{bmatrix} P_i^{std} \\ P_i^{sha} \end{bmatrix} = F(M_i, S_i, R_i, U_i)$$

where F is a known function, M_i is the (observed) treatment status from the ongoing pricing experiments, S_i is an indicator for whether or not the observation is in the ongoing pricing experiment, R_i is the (observed) treatment status from the LDB experiment, and U_i are the (observed) determinants of price that are not explicitly randomized. Moreover, $(M_i, S_i) \perp R_i$ and

$$(M_i, S_i, R_i) \perp U_i, \theta_i^{std}, \theta_i^{sha}, V_i^{std}, V_i^{sha}.$$

Given these assumptions, we can prove the following result:

Lemma 1. *Suppose the data are generated according to 3–5. Let* $P_i = (P_i^{std}, P_i^{sha})'$. *Then*

$$\begin{aligned} \begin{bmatrix} \beta^{std,std}(u,t) \\ \beta^{std,sha}(u,t) \end{bmatrix} &= \mathrm{Var}(P_i|U_i = u, \theta_i = t)^{-1}\mathrm{Cov}(P_i, Y_i^{*,std}(P_i)|U_i = u, \theta_i = t) \\ &= \mathrm{Var}(\hat{P}_i|U_i = u, \theta_i = t)^{-1}\mathrm{Cov}(\hat{P}_i, Y_i^{std}|U_i = u, \theta_i = t) \end{aligned}$$

20. In particular, we restrict our attention to sessions where a price was within 10 cents of the nearest dollar.

and

$$\begin{bmatrix}\beta^{sha,std}(u,t)\\ \beta^{sha,sha}(u,t)\end{bmatrix} = \mathrm{Var}(P_i|U_i=u,\theta_i=t)^{-1}\,\mathrm{Cov}(P_i, Y_i^{*,sha}(P_i)|U_i=u,\theta_i=t)$$
$$= \mathrm{Var}(\hat{P}_i|U_i=u,\theta_i=t)^{-1}\,\mathrm{Cov}(\hat{P}_i, Y_i^{sha}|U_i=u,\theta_i=t)$$

Proof. Assumption 4′ ensures that fixing the values of U_i and θ_i, price variation is exogenous. As a result, subsetting in this way, the conditional expectation function with respect to price is identical to the expected potential outcome function. Assumption 5 ensures that subsetting to observations with fixed U_i and θ_i, the linear model is correctly specified, so the result follows from standard propoerties of OLS under correct specification. The proof of the second equality follows identically. □

Corollary 4.

$$\mathbb{E}\begin{bmatrix}\beta^{std,std}(U_i,\theta_i)\\ \beta^{std,sha}(U_i,\theta_i)\end{bmatrix} = \mathbb{E}[\mathrm{Var}(\hat{P}_i \mid U_i, S_i=1)^{-1}(Y_i^{std}(\hat{P}_i - \mathbb{E}[\hat{P}_i \mid U_i, S_i=1])) \mid S_i=1]$$
$$\mathbb{E}\begin{bmatrix}\beta^{sha,std}(U_i)\\ \beta^{sha,sha}(U_i)\end{bmatrix} = \mathbb{E}[\mathrm{Var}(\hat{P}_i \mid U_i, S_i=1)^{-1}(Y_i^{sha}(\hat{P}_i - \mathbb{E}[\hat{P}_i \mid U_i, S_i=1])) \mid S_i=1]$$

Proof. Independence implies that conditioning on S_i has no effect on the resulting expectation. Lemma 1 implies that conditional on $U_i=u_i, \theta_i=t_i$, the two expectations are identical. Moreover, the conditional variance and expectation of P_i given U_i does not depend on θ_i. The result then follows by the law of iterated expectations. □

The above result does not directly allow us to compute the slopes of the latent demand curves because they require us to use $\hat{P}_i$, which is unobserved and depends on the inattention parameter θ_i. In practice, we simply replace $\hat{P}_i$ with P_i, which we show in Appendix C is unlikely to substantially alter our results.[21] Because we know Lyft's pricing algorithm, the quantities $\mathrm{Var}(P_i \mid U_i)$ and $\mathbb{E}[P_i \mid U_i]$ can be directly computed for each observation. We can thus form the sample analog of the expectations on the RHS above.

As in our conceptual framework, we now show how to compare these identified latent demand slopes to the size of the demand discontinuity near integral prices to recover the inattention parameter. Here, we focus on the variation specifically generated by our LDB experiment. Not every price is affected by this experiment, and in some cases, one price but not the other is affected. To help distinguish between the cases, we define $P_i^{cf} = (P_i^{std,cf}, P_i^{sha,cf})'$ where $P_i^{std,cf}$ is the price of Standard if the observation had been in the control group of the LDB experiment while $P_i^{sha,cf}$ is the price of Shared if the observation had been in the control group. Clearly, these prices are deterministic functions of M_i and U_i and hence independent of which treatment group in the LDB experiment the observation actually was from.

The structure of the LDB experiment ensures that for any fixed value of P_i^{cf}, there are at most two possible values of P_i^{std} and at most two values of P_i^{sha}. We therefore define the variable $D_i = (D_i^{std}, D_i^{sha})'$ where D_i^{std} is a dummy variable for if Standard price was changed by the

21. Intuitively, this is because the variation we are using to estimate the slope of the latent demand curve spans multiple dollar amounts. As a result, distortions to perceived price due to left-digit bias are small relative to the underlying variation in prices. Moreover, insofar as the pricing algorithm up until our experiment was largely agnostic about cents amounts, we expect that many of the effects of left-digit bias will cancel out from observation to observation.

LDB experiment, *i.e.* $P_i^{std} \neq P_i^{std,cf}$, and D_i^{sha} is a dummy variable for if Shared price was changed by the LDB experiment, *i.e.* $P_i^{sha} \neq P_i^{sha,cf}$.

The following Lemma combines the idea behind Lemma 1 with the ideas behind Propositions 2 and 4 and will be useful for understanding what the effects observed in the LDB experiment correspond to:

Lemma 2. *Suppose the data are generated according to Assumptions 3–5. Suppose P_i^{cf} is such that only P_i^{std} is affected by the LDB experiment.*

$$\frac{\mathrm{Cov}(Y_i^j, D_i^{std} \mid P_i^{cf}, \theta_i)}{\mathrm{Var}(D_i^{std} \mid P_i^{cf}, \theta_i)} = \Delta \hat{P}_i^{std}(\theta_i^{std}, P_i^{std,cf})\mathbb{E}[\beta^{j,std}(U_i, \theta_i) \mid P_i^{cf}]$$

where $\Delta \hat{P}_i^{std}(\theta_i, P_i^{cf})$ is the amount by which the effect of the LDB experiment affects Standard price and $j = std, sha$ indexes the product whose demand is being measured.

Similarly, if only P_i^{sha} is affected by the LDB experiment,

$$\frac{\mathrm{Cov}(Y_i^j, D_i^{sha} \mid P_i^{cf}, \theta_i)}{\mathrm{Var}(D_i^{sha} \mid P_i^{cf}, \theta_i)} = \Delta \hat{P}_i^{sha}(\theta_i^{sha}, P_i^{sha,cf})\mathbb{E}[\beta^{j,sha}(U_i, \theta_i) \mid P_i^{cf}]$$

where $\Delta \hat{P}_i^{sha}(\theta_i, P_i^{cf})$ is defined analogously.

Finally, if P_i^{cf} is such that both prices are affected, then

$$\mathrm{Var}(D_i \mid P_i^{cf}, \theta_i)^{-1}\mathrm{Cov}(Y_i, D_i \mid P_i^{cf}, \theta_i) = \begin{bmatrix} \Delta \hat{P}_i^{std}(\theta_i^{std}, P_i^{std,cf})\mathbb{E}[\beta^{j,std}(U_i, \theta_i) \mid P_i^{cf}] \\ \Delta \hat{P}_i^{sha}(\theta_i^{sha}, P_i^{sha,cf})\mathbb{E}[\beta^{j,sha}(U_i, \theta_i) \mid P_i^{cf}] \end{bmatrix}$$

Proof. Conditional on θ_i, D_i is collinear with perceived prices and is a function of R_i. Because after conditioning on P_i^{cf}, the only residual variation in perceived prices is due to R_i, which was randomized, the conditional expectation function given D_i is the expected potential outcome. The result then follows from the fact that OLS is correctly specified given Assumption 5. □

We now explicitly solve for what $\Delta \hat{P}_i^{std}$ and $\Delta \hat{P}_i^{sha}$ are. For Standard, either $¢P_i^{std,cf} \in [0.0.09]$ and

$$\begin{aligned} \Delta \hat{P}_i^{std,cf} &= (1-\theta^{std})¢P_i^{std,cf} + 1 - (1-\theta_i^{std})(0.99 - ¢P_i^{std,cf}) \\ &= \underbrace{0.01 + ¢P_i^{std,cf}}_{1-\lambda_i^{std}} + \theta\underbrace{(0.99 - ¢P_i^{std,cf})}_{\lambda_i^{std}} \end{aligned}$$

or otherwise, D_i^{std} is uniformly 0. For Shared prices, either $¢P_i^{sha,cf} \in [0.0.09]$ in which case

$$\begin{aligned} \Delta \hat{P}_i^{sha,cf} &= (1-\theta^{sha})¢P_i^{sha,cf} + 1 - (1-\theta_i^{sha})(0.99 - ¢P_i^{sha,cf}) \\ &= \underbrace{0.01 + ¢P_i^{sha,cf}}_{1-\lambda_i^{sha}} + \theta\underbrace{(0.99 - ¢P_i^{sha,cf})}_{\lambda_i^{sha}} \end{aligned}$$

or $\text{¢}P_i^{sha,cf} \in [0.9, 0.99]$ in which case

$$\begin{aligned}\Delta \hat{P}_i^{sha,cf} &= (1-\theta)\text{¢}P_i^{sha,cf} - 1 \\ &= \underbrace{\text{¢}P_i^{sha,cf} - 1}_{-(1+\lambda_i)} + \theta \underbrace{(-\text{¢}P_i^{sha,cf})}_{-\lambda_i}\end{aligned}$$

or otherwise, D_i^{sha} is uniformly constant.

Now define the vector $T_i = (T_i^{std}, T_i^{sha})'$ such that

$$T_i^{std} = \begin{cases} \dfrac{D_i^{std} - \mathbb{E}[D_i^{std} \mid P_i^{cf}]}{\mathrm{Var}(D_i^{std}|P_i^{cf})} & \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0, \mathrm{Var}(D_i^{sha} \mid P_i^{cf}) = 0 \\ [\mathrm{Var}(D_i \mid P_i^{cf})^{-1}(D_i - \mathbb{E}[D_i^{std} \mid P_i^{cf}])]^{std} & \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0, \mathrm{Var}(D_i^{sha} \mid P_i^{cf}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$T_i^{sha} = \begin{cases} \dfrac{D_i^{sha} - \mathbb{E}[D_i^{sha} \mid P_i^{cf}]}{\mathrm{Var}(D_i^{sha}|P_i^{cf})} & \mathrm{Var}(D_i^{sha} \mid P_i^{cf}) > 0, \mathrm{Var}(D_i^{std} \mid P_i^{cf}) = 0 \\ [\mathrm{Var}(D_i \mid P_i^{cf})^{-1}(D_i - \mathbb{E}[D_i^{sha} \mid P_i^{cf}])]^{sha} & \mathrm{Var}(D_i^{sha} \mid P_i^{cf}) > 0, \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where above, for a vector v, v^k here means the component of the vector v corresponding to the product k. Again, because we know algorithmically how prices are set, the conditional expectations and variances of D_i can be directly computed. The following result shows how T_i can be used to construct moments that can help us identify θ_i:

Corollary 5.

$$\mathbb{E}[T_i^{std} Y_i^j \mid \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0] = \mathbb{E}[\beta^{j,std}(U_i, \theta_i) \mid \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0] \left(\overline{\lambda^{j,std}} \cdot \overline{\theta^{j,std}} + 1 - \overline{\lambda^{j,std}}\right)$$

$$\mathbb{E}[T_i^{sha} Y_i^j \mid \mathrm{Var}(D_i^{sha} \mid P_i^{cf}) > 0] = \mathbb{E}[\beta^{j,sha}(U_i, \theta_i) \mid \mathrm{Var}(D_i^{sha} \mid P_i^{cf}) > 0] \left(\overline{\lambda^{j,sha}} \cdot \overline{\theta^{j,sha}} + 1 - \overline{\lambda^{j,sha}}\right)$$

where

$$\overline{\lambda^{j,std}} = \frac{\mathbb{E}[\beta^{j,std}(U_i, \theta_i)\lambda_i^{std} \mid \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0]}{\mathbb{E}[\beta^{j,std}(U_i, \theta_i) \mid \mathrm{Var}(D_i^{std} \mid P_i^{cf}) > 0]}$$

$$\overline{\theta^{j,std}} = \frac{\mathbb{E}[\beta^{j,std}(U_i, \theta_i)\lambda_i^{std}\theta_i^{std}|\mathrm{Var}(D_i^{std}|P_i^{cf}) > 0]}{\mathbb{E}[\beta^{j,std}(U_i, \theta_i)\lambda_i^{std}|\mathrm{Var}(D_i^{std}|P_i^{cf}) > 0]}$$

$$\overline{\lambda^{j,sha}} = \frac{\mathbb{E}[\beta^{j,sha}(U_i, \theta_i)\lambda_i^{sha}|\mathrm{Var}(D_i^{sha}|P_i^{cf}) > 0]}{\mathbb{E}[\beta^{j,sha}(U_i, \theta_i)|\mathrm{Var}(D_i^{sha}|P_i^{cf}) > 0]}$$

$$\overline{\theta^{j,sha}} = \frac{\mathbb{E}[\beta^{j,sha}(U_i, \theta_i)\lambda_i^{sha}\theta_i^{sha}|\mathrm{Var}(D_i^{sha}|P_i^{cf}) > 0]}{\mathbb{E}[\beta^{j,sha}(U_i, \theta_i)\lambda_i^{sha}|\mathrm{Var}(D_i^{sha}|P_i^{cf}) > 0]}$$

Proof. This follows directly from Lemma 2 and the law of iterated expectations. □

Intuitively, we would like to be able to combine Corollaries 4 and 5 to identify inattention parameters by "dividing" the LHS of Corollary 5 by the RHS of 4 and solving for the $\overline{\theta}$'s. We need to make a number of assumptions, to justify doing so:

Assumption 6.

$$\mathbb{E}[\beta^{j,k}(U_i,\theta_i)\,|\,\mathrm{Var}(D_i^k\,|\,P_i^{cf})>0]=\mathbb{E}[\beta^{j,k}(U_i,\theta_i)]$$

where $j=std, sha$ and $k=std, sha$. Moreover,

$$\mathrm{Cov}(\beta^{j,std}(U_i,\theta_i),\lambda_i^k)=0$$

Remark 3. Both assumptions essentially require that there is no systematic relationship between the slope of the latent demand curve and the precise cents amount of the control price. Substantial violations of these assumptions require that within-dollar, the conditional average slope of the latent demand curve covaries substantially with the cents amount. Looking at the relative smoothness of the measured elasticity function in Figures 6 and 7, this would require that price sensitivity is a highly non-monotone function of control price, which we view as implausible.

Given Assumption 6, $\overline{\lambda^{j,k}}=\mathbb{E}[\lambda_i^k\,|\,\mathrm{Var}(D_i^k>0)]\equiv\overline{\lambda^j}$, which is therefore directly observable. Moreover, combining this assumption with Corollaries 4 and 5 implies that

$$\frac{\mathbb{E}[T_i^kY_i^j\,|\,\mathrm{Var}(D_i^k>0)]}{\mathbb{E}[\mathrm{Var}(P_i\,|\,U_i)^{-1}(Y_i^j\cdot(P_i-\mathbb{E}[P_i\,|\,U_i]))\,|\,S_i=1]^k}=\overline{\lambda^j}\cdot\overline{\theta^{j,k}}+1-\overline{\lambda^j}$$

We can clearly solve this equation for $\overline{\theta^{j,k}}$, which implies that using the assumptions we have made so far (which we argue are all either mild or implied by our experimental randomization), we can identify the parameters $\overline{\theta^{j,k}}$, which we can think of as a slope-weighted average inattention in the population.[22] It can be interpreted as the local average inattention parameter subsetting to individuals who are just at the margin of changing their consumption of good j when the price of good k changes.

In principle, even if the assumptions we have been making so far are accurate, $\overline{\theta^{j,k}}$ may take on a different value for each of the $(4=2\times 2)$ choices of j and k. In principle, we could therefore estimate each $\overline{\theta^{j,k}}$ separately by plugging in the relevant moments. However, putting these moments into a GMM framework allows us to specify tests that tell us whether a more parsimonious model of inattention is also consistent with the data:

Assumption 7.

$$\mathrm{Cov}(\theta_i,\beta^{j,k}(U_i,\theta_i))=0$$

Assumption 8. $\theta_i^{std}=\theta_i^{sha}$.

Under these assumptions, there is some common $\overline{\theta}=\overline{\theta^{j,k}}$ for all j,k. We can thus describe our parameter vector as consisting of $\overline{\beta^{j,k}}=\mathbb{E}[\beta^{j,k}(U_i,\theta_i)]$ and $\overline{\theta}$. These parameters satisfy the following moment conditions:

$$0=\mathbb{E}[S_i([\mathrm{Var}(P_i|U_i)^{-1}(Y_i^j\cdot(P_i-\mathbb{E}[P_i|U_i]))]^j-\overline{\beta^{j,k}})]$$
$$0=\mathbb{E}[\mathbb{1}\{\mathrm{Var}(D_i^k|P_i^{cf})>0\}(T_i^k(Y_i^j-\mathbb{E}[Y_i^j])-(\lambda_i^j\bar{\theta}+\mathrm{sign}(\lambda_i^j)|1-\lambda_i^j|)\,\mathrm{sign}(\lambda_i^j))]$$

Including $\mathbb{E}[Y_i^j]$ in the second set of moment conditions does not change its solution, since it is multiplied by a mean 0 constant. However, it does help reduce the variance of our estimator, which is especially important because while our slope parameters are precisely estimated, the

22. Moreover, monotone comparative statics results imply that the weights are all non-negative.

estimates of the discontinuity due to the experiment are noisier. Moreover, asymptotic standard errors will be unaffected if we replace $\mathbb{E}[Y_i^j]$ with its sample mean, which is what we do in practice.

We thus estimate our single inattention parameter by minimizing this GMM objective with respect to our five parameters and test our overidentifying restrictions using the corresponding Sargan-Hansen test. This can be thought of as a joint test of the inattention model and of the assumption that θ_i is relatively homogeneous conditional on the inattention model being true. We are unable to reject the null hypothesis that our model is correctly specified.

E.1. *Justification for using actual prices*

Similar to what we do for the observational data, in this subsection, we justify our choice to take slopes estimated off of the multiplier-based pricing variation as being the same as slopes of the latent demand curve. In the discussion following Corollary 4, we noted that in practice, we proceeded by replacing $\hat{P}_i$ with P_i. Intuitively, this strategy should be unlikely to substantially alter results when we are considering pricing variation that spans multiple dollar amounts. First, inattention can at most change perceived prices within the dollar amount, so any errors due to replacing $\hat{P}_i$ with P_i can be expected to be small relative to the variance of the distribution of prices. Second, to the extent that the cents amounts in prices are random, we may even expect these errors to cancel out. Formalizing this intuition appears difficult, so we instead opt here to demonstrate that it is empirically justified. Specifically, recall that our estimates for the slopes of the demand curve aggregate the OLS slopes from the regression of Y_i on P_i conditional on U_i and θ_i.

Thus, fixing θ_i and U_i, we ideally wish to estimate a regression of the form

$$Y_i = \beta_0 + \hat{P}_i'\beta + \varepsilon_i, \quad \mathbb{E}[\varepsilon_i \hat{P}_i] = 0, \quad \hat{P}_i = P_i + \delta_i \tag{13}$$

where $\delta_i = \theta_i \cdot (P_i - \lfloor P_i \rfloor)$. Moreover, because P_i spans multiple dollar amounts, $\mathrm{Cov}(P_i, \delta_i) \approx 0$. Instead of the ideal regression, we instead regress Y_i on P_i, which therefore yields the coefficient

$$\begin{aligned}
\beta^* &= \mathbb{E}[(P_i - \mathbb{E}[P_i]) \cdot P_i^T]^{-1}\mathbb{E}[(P_i - \mathbb{E}[P_i]) \cdot \hat{P}_i^T]\beta \\
&= \mathbb{E}[(P_i - \mathbb{E}[P_i]) \cdot (\hat{P}_i^T - \delta_i)]^{-1}\mathbb{E}[(P_i - \mathbb{E}[P_i]) \cdot \hat{P}_i^T]\beta \\
&\approx (I + \mathbb{E}[(P_i - \mathbb{E}[P_i]) \cdot \hat{P}_i^T]^{-1}\mathrm{Cov}(P_i, \delta_i)) \cdot \beta
\end{aligned} \tag{14}$$

This implies that an upper bound on the relative bias is roughly

$$\|\mathrm{Cov}(P_i, \hat{P}_i)^{-1}\mathrm{Cov}(P_i, P_i - \lfloor P_i \rfloor) \cdot \theta\| \tag{15}$$

Thus, as long as $\max_\theta \|\mathrm{Cov}(P_i, \$P_i + (1-\theta)¢P_i)^{-1}\mathrm{Cov}(P_i, \theta \cdot (P_i - \lfloor P_i \rfloor))\| \ll 1$, the potential bias from ignoring inattention when looking at the exogenous multiplier variation is minimal. Roughly speaking, the way that the price multipliers are incorporated into a final price is via the formulae

$$\begin{aligned}
P^{std} &= P^{std,cf} \cdot M^{std} \\
P^{sha} &= \min(P^{std} \cdot 0.9, P^{sha,cf} \cdot M^{sha})
\end{aligned} \tag{16}$$

where $P^{j,cf}$ are counterfactual prices (*i.e.* prices in the absence of the multiplier). Based on this mapping from counterfactual prices and multipliers to realized prices, and our knowledge of the distribution of multipliers, we run the following procedure

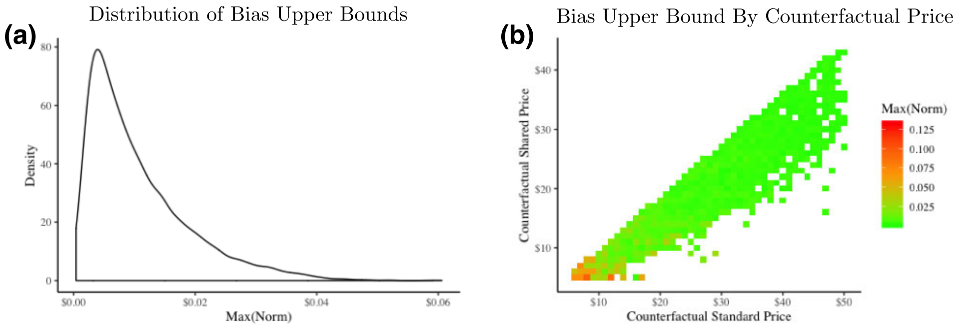


FIGURE A.5

Worst-case bias in structural estimates from experiment

Notes: This figure estimates the worst-case bias described by equation (15) in Appendix E.1. Panel (a) plots the distribution of worst-case biases from using realized instead of perceived prices to estimate elasticity of latent demand as described in equation (15) across different dollar amounts for Standard and Shared price. Panel (b) shows how this worst-case bias varies by price.

TABLE A1

Inattention parameter from experimental data

	All sessions	Counterfactual prices >= 7.5
Std. to Std.	−0.036	−0.029
	(0.0009)	(0.0008)
Std. to Sha.	0.012	0.012
	(0.0005)	(0.0004)
Sha. to Std.	0.011	0.0088
	(0.001)	(0.0001)
Sha. to Sha.	−0.024	−0.020
	(0.0007)	(0.0007)
Inattention (θ)	0.45	0.48
	(0.02)	(0.03)

Notes: This table replicates the results in the first column of Table 8 using only sessions where the counterfactual price is sufficiently large. The first four rows display estimated derivatives of the latent demand curve while the last column displays estimated inattention. A derivative labeled X to Y means the derivative of Y demand with respect to exogenous changes in X price. Standard errors are in parentheses. The first column above is just a re-print of the results from the first column of Table 8. The second column replicates the methodology used to compute these estimates, but subsets to sessions where the counterfactual price of Standard and Shared are both >= 7.5.

1. Draw $P^{std,cf}$, $P^{sha,cf}$ from the distribution of sessions
2. Simulate 10,000 bootstrap samples of P using the known distributions of M^{std}, M^{sha}
3. Compute $\max_\theta \|\mathrm{Cov}(P, \hat{P}_\theta)^{-1}\mathrm{Cov}(P, \theta \cdot (P - \lfloor P \rfloor))\|$

In Figure A5, we plot the distribution of $\max_\theta \|\mathrm{Cov}(P, \hat{P}_\theta)^{-1}\mathrm{Cov}(P, \theta \cdot (P - \lfloor P \rfloor))\|$ produced with this procedure and how it varies with counterfactual prices. We find that most of the time, this upper bound is low, and the only times it is not low are when prices are cheap.[23] As an

23. These worst-case bounds are likely to be very conservative. If we take $\theta = 0.450$ to be the true level of inattention (instead of maximizing over θ), the worst-case bound instead is only 0.06, even for the worst prices. Additionally, even if the worst case is large for a fixed price, there is likely to be additional canceling out among the sessions that are badly biased.

additional robustness check, we re-estimate inattention, now subsetting to sessions with counterfactual prices above \$7.5.[24] Results are displayed in Table A1 and find that our estimated inattention parameter does not change by much.

Acknowledgments. We wish to thank the Editor, Nicola Gennaioli, and three anonymous reporters for excellent remarks that improved the study. Stefano DellaVigna, Sam Hirshman, Avner Strulov-Shlain, and Justin Sydnor also provided quite helpful comments along with seminar participants at Berkeley, Datacolada, University of Arizona, University of Southern California, and West Point Military Academy.

Data Availability Statement

The primary data for this project are confidential and were provided by Lyft under a Data Use Agreement (DUA) that prohibits making the data publicly available. Academic researchers interested in working with Lyft data can contact Klaus Ashorn (klausa@lyft.com). Requests for data access will be reviewed by Lyft on a case-by-case basis. Replication code (List *et al*., 2022) for this project, along with simulated data on which the replication code runs available on Zenodo at https://doi.org/10.5281/zenodo.7379460.

REFERENCES

ANDERSON, E. and SIMESTER, D. (2003), "Effects of \$9 Price Endings on Retail Sales: Evidence from Field Experiments", *Quantitative Marketing and Economics*, **1**: 93–110.
ANGRIST, J. D. and IMBENS, G. W. (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, **62**, 467–475.
ANGRIST, J. D., GRADDY, K. and IMBENS, G. W. (2000), "The Interpretation of Instrumental Variables Estimators in Simultaneous Equation Models with an Application to the Demand for Fish", *Review of Economic Studies*, **67**, 499–527.
ARMSTRONG, S. and VICKERS, J. (2015), "Which Demand Systems Can be Generated by Discrete Choice?", *Journal of Economic Theory*, **158**, 293–307.
ASHTON, L. (2014), "Left-Digit Bias and Inattention in Retail Prices: Evidence from a Field Experiment", *SSRN Electronic Journal*. doi:10.2139/ssrn.2538816
BASU, K. (1997), "Why Are So Many Goods Priced to End in Nine? And Why This Practice Hurts the Producers", *Economics Letters*, **54**, 41–44.
BASU, K. (2006), "Consumer Cognition and Pricing in the Nines in Oligopolistic Markets", *Journal of Economics & Management Strategy*, **15**, 125–141.
BECKER, G. S. (1962), "Irrational Behavior and Economic Theory", *Journal of Political Economy*, **70**, 1–13.
BLOOM, N. and VAN REENEN, J. (2007), "Measuring and Explaining Management Practices Across Firms and Countries", *The Quarterly Journal of Economics*, **122**, 1351–1408.
BUSSE, M. R., KNITTEL, C. R. and ZETTELMEYER, F. (2013), "Are Consumers Myopic? Evidence from New and Used Car Purchases", *American Economic Review*, **103**, 220–256.
CARVER, J. R. and PADGETT, D. T. (2012), "Product Category Pricing and Future Price Attractiveness: 99-Ending Pricing in a Memory-Based Context", *Journal of Retailing*, **88**, 497–511.
CHETTY, R., LOONEY, A. and KROFT, K. (2009), "Salience and Taxation: Theory and Evidence", *American Economic Review*, **99**, 1145–1177.
COHEN, P., HAHN, R. and HALL, J., *et al.* (2016), "Using Big Data to Estimate Consumer Surplus: The Case of Uber" (NBER Working Paper).
CONLIN, M., O'DONOGHUE, T. and VOGELSANG, T. J. (2007), "Projection Bias in Catalog Orders", *American Economic Review*, **97**, 1217–1249.
CONLON, C. T. and RAO, N. L. (2020), "Discrete Prices and the Incidence and Efficiency of Excise Taxes", *American Economic Journal: Economic Policy*, **12**, 111–143.
DELLAVIGNA, S. (2009), "Psychology and Economics: Evidence from the Field", *Journal of Economic Literature*, **44**: 315-372.
DELLAVIGNA, S. and GENTZKOW, M. (2019), "Uniform Pricing in US Retail Chains", *SSRN Electronic Journal*. doi:10.2139/ssrn.3367978
DELLAVIGNA, S., LIST, J. A. and MALMENDIER, U. (2012), "Testing for Altruism and Social Pressure in Charitable Giving", *The Quarterly journal of Economics*, **127**, 1–56.
DELLAVIGNA, S., LIST, J. A. and MALMENDIER, U., *et al.* (2016a), "Voting to Tell Others", *The Review of Economic Studies*, **84**, 143–181.

24. Looking at Figure A5(b), the value of Max(Norm) appears to be relatively small above \$8. We include the 50 cents below \$8 to include rounding near \$8.

DELLAVIGNA, S., LIST, J. A. and MALMENDIER, U., *et al.* (2016b), "Estimating Social Preferences and Gift Exchange at Work" (NBER Working Paper).

DELLAVIGNA, S. and MALMENDIER, U. (2004), "Contract Design and Self-Control: Theory and Evidence", *The Quarterly Journal of Economics*, **119**, 353–402.

DUBE, A., MANNING, A. and NAIDU, S. (2020), "Monopsony, Misoptimization, and Round Number Bunching in the Wage Distribution" (NBER Working Paper w24991).

ENGLMAIER, F., SCHMOLLER, A. and STOWASSER, T. (2017), "Price Discontinuities in an Online Market for Used Cars", *Management Science*, **64**: 2754-2766.

GERRITSEN, A. (2015), "Optimal Taxation When People Do Not Maximize Well-Being", *Journal of Public Economics*, **144**, 122–139.

GOLDFARB, A. and XIAO, M. (2011), "Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets", *American Economic Review*, **101**, 3130–3161.

GYÖRFI, L., KOHLER, M. and KRZYŻAK, A., *et al.* (2002), *A Distribution-Free Theory of Nonparametric Regression* (Springer Series in Statistics).

HANNA, R., MULLAINATHAN, S. and SCHWARTZSTEIN, J. (2014), "Learning Through Noticing: Theory and Evidence from a Field Experiment", *The Quarterly Journal of Economics*, **129**, 1311–1353.

HARRISON, G. W. and LIST, J. A. (2004), "Field Experiments", *Journal of Economic Literature*, **42**, 1009–1055.

HILGER, N. (2018), "Heuristic Thinking in the Market for Online Subscriptions", Available at SSRN 3296698.

HORTAÇSU, A. and PULLER, S. L. (2008), "Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market", *RAND Journal of Economics*, **39**, 86–114.

JIN, G. Z., KATO, A. and LIST, J. A. (2010), "That's News To Me! Information Revelation In Professional Certification Markets", *Economic Inquiry, Western Economic Association International*, **48**, 104–122.

LACETERA, N., POPE, D. G. and SYDNOR, J. R. (2012), "Estimating the Effect of Salience in Wholesale and Retail Car Markets", *American Economic Review*, **103**, 575–579.

LAIBSON, D., REPETTO, A. and TOBACMAN, J. (2007), "Estimating Discount Functions with Consumption Choices over the Lifecycle" (NBER Working Paper).

LIST, J. A. (2003), "Does Market Experience Eliminate Market Anomalies?", *The Quarterly Journal of Economics*, **118**, 41–71.

LIST, J. A. (2004), "Testing Neoclassical Competitive Theory in Multilateral Decentralized Markets", *Journal of Political Economy*, **112**, 1131–1156.

LIST, J. A., MUIR, I. and POPE, D. G., *et al.* (2022), "Replication data for: Left-Digit Bias at Lyft", Review of Economic Studies, Zenodo.

MASSEY, C. and THALER, R. H. (2013), "The Loser's Curse: Decision Making and Market Efficiency in the National Football League Draft", *Management Science*, **59**, 1479–1495.

MENG, C. (2020), *Prospect Theory in the Housing Market* (Doctoral dissertation, University of Cambridge).

MURPHY, K. (2012), *Machine Learning: A Probabalistic Perspective* (The MIT Press).

NEGI, A. and WOOLDRIDGE, J. M. (2019), "Revisiting Regression Adjustment in Experiments with Heterogeneous Treatment Effects", *Econometric Reviews*.

REPETTO, L. and SOLÍS, A. (2020), "The Price of Inattention: Evidence from the Swedish Housing Market", *Journal of the European Economic Association*, **18**, 3261–3304.

SIMON, H. A. (1957), *Models of Man* (New York: John Wiley).

STIVING, M. (2000), "Price-Endings when Prices Signal Quality", *Management Science*, **46**, 1617–1629.

STIVING, M. and WINER, R. S. (1997), "An Empirical analysis of Price Endings with Scanner Data", *Journal of Consumer Research*, **24**, 57–67.

STRULOV-SHLAIN, A. (2021), "More Than a Penny's Worth: Left-Digit Bias and Firm Pricing", *SSRN Electronic Journal*. doi:10.2139/ssrn.3413019