More than a Penny's Worth: Left-Digit Bias and Firm Pricing

Avner S. Shlain*

University of California, Berkeley

September 14, 2018

Abstract

Why do so many prices end with 99 cents? Firms arguably price at 99-ending prices because of left-digit bias, the tendency of consumers to perceive a $4.99 as much lower than $5.00. Using retail scanner data on thousands of products and dozens of retailers, I provide reduced-form support for this explanation. I then structurally estimate the magnitude of left-digit bias, and find that consumers respond to a 99-ending price as if it were 15-25 cents lower. Next, I analyze how should firms respond to left-digit biased demand. I solve and estimate a model that makes three key predictions: (1) prices should bunch at 99-ending prices; (2) there should be ranges of missing prices with low price-endings; (3) these ranges of missing prices should increase with the dollar digit. Qualitatively, these predictions hold. Firms respond to the bias with high shares of 99s and missing low-ending prices. Quantitatively, however, firms price *as if* the bias were much smaller and demand were more elastic. I estimate that the retailer is forgoing 1-3 percents of potential gross profits due to this misperception.

1 Introduction

Retail prices commonly end with 99 cents (Ginzberg (1936); Conlon and Rao (2016); Stiving and Winer (1997))[1]. A leading explanation for this pricing behavior is that consumers are left-digit biased, meaning they ignore the cents component of the price (referred to as "level effects", e.g. Bizer and Schindler (2005)). An alternative explanation, "image effects", is that consumers learn about a product's higher quality or value from its price-ending (Stiving (2000)).

As common as 99-ending pricing is, we do not have a model of left-digit biased demand that is supported by the data and which can be integrated in optimal pricing models. As a consequence, it is difficult to evaluate the impact of this pricing behavior on consumers and firms. What level, if any, of left-digit bias do consumers exhibit? In turn, how should firms respond to the bias? Is the observed prevalence of 99-ending prices optimal, given the observed degree of left-digit bias by consumers? If not, how sizable are the profit losses?

To answer these questions, I provide a portable model, exmaine whether it is supported by the data, and estimate its parameters. I integrate the structure of left-digit biased consumers of Lacetera, Pope and Sydnor (2012) with a model of profit-maximizing firms to derive a set of predictions on pricing and profits. First, I provide reduced-form evidence for left-digit bias in consumer demand, and structurally estimate the degree of the bias on thousands of products in scanner data. I find that consumers are biased, to the extent of treating a 99 ending price as if it were 15-25 cents lower. Next, I estimate retailer pricing behavior. I find that firms pricing is aligned with the model's main predictions, but, they seem to underestimate the magnitude of the bias significantly. From the firm's perspective, I estimate that they respond as if a 99-ending price is treated by consumers as being only 0.5-3 cents lower. Finally, because of these discrepancies, I calculate economically significant implications on firm profits.

The first step in the paper is to create a framework to understand the phenomenon. As in Chetty et al. (2009); DellaVigna (2009); Lacetera et al. (2012), I model left-digit bias as if consumers ignore a θ share of a component of the price – the lower digits. Focusing in this paper on single dollar-digit items, this makes a price p to be perceived as $\hat{p} = (1 - \theta) p + \theta \lfloor p \rfloor$ (a mix with weight θ on the dollar digits, i.e. the price with no cents, and $1 - \theta$ on the exact price). For example, with a bias of 0.15, a consumer perceives \$4.00 as \$4.00, but a \$3.99 as \$3.85. The bias causes aggregate demand to drop when the first digit changes, e.g. from 3.99 to 4.00, and also dampens the price sensitivity to within-dollar price changes, e.g. from \$4.00 to \$4.01.

How should a firm respond to consumers who perceive prices this way? Optimal pricing exhibits bunching at 99-ending prices and missing low-ending prices. As left-digit bias increases, as demand becomes less elastic, or for larger dollar-digits, the missing prices regions increase.

Next, I turn to explore whether the bias is evident on the demand side. A vast literature has explored the issue, with positive but imprecise results (see Stiving and Winer (1997); Macé (2012) for reviews and results). First, proper data cleaning and big scanner data allow to give finer and cleaner corroborating evidence for the bias and rule out other explanations; Second, unlike existing literature, I estimate the bias such that we can incorporate it into the firms problem. To estimate demand, I use a sample of 1710 popular products in 248 stores of a single US retailer over 3.5 years, and 12 products in AC Nielsen RMS data across more than 60 chains and 11,000 stores, over 9 years. The key idea is that if the model is correct, controlling for demand shifters, we should observe drops in demand crossing the dollar digits. This holds even though firms are strategic and price-endings are endogenous, as long as there is some variation in price-endings.

---

[1]See also Aalto-Setala (2005); Anderson et al. (2015); Anderson and Simester (2003); Ashton (2014); Ater and Gerlitz (2017); El Sehity et al. (2005); Kalyanam and Shively (1998); Levy et al. (2011); Macé (2012); Schindler and Kibarian (1996); Shlain (2018); Snir et al. (2017)

Figure A provides an example of a demand curve for a single product, Tillamook medium cheddar cheese slices. The horizontal axis is the price and vertical axis is log-quantity sold, residualized from various controls. Each point is an estimate of the average demand at that price, with standard errors clustered at the store level. Colors represent the dollar digit. Details regarding data and estimation are described in the body of the paper. The figure shows that demand is flatter within dollars than across all prices (solid lines versus dashed line), with drops between dollars. These are the patterns expected if consumers exhibit some form of left-digit bias.

Figure A: Demand curve of a single product, exhibiting drops at dollar digits
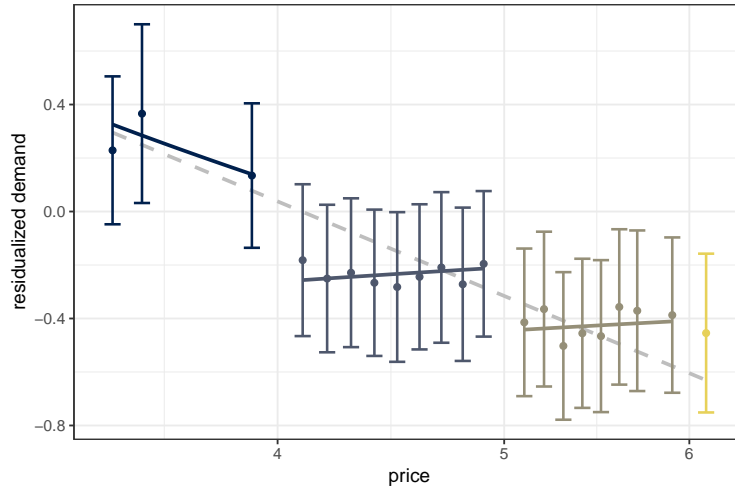
As a one-sided visual test of the model, I focus on the structure of the unexplained component of demand. I residualize quantities purchased from the overall price sensitivity (e.g., with constant elasticity), and other product level controls such as seasonality, substitution with similar products, and promotions; then, I consider the *residuals by price*. Relative to a smooth decreasing demand, the unexplained component in demand should be lower than the best fit for low price endings, and higher for high price endings[2]. So, residuals should look like a sawtooth pattern - increasing within each dollar digit and dropping sharply across the dollars. The main advantage of this procedure is that it allows to aggregate over different products (with different elasticities and average prices).

Several empirical challenges exist when trying to get clean evidence of residuals by price from retail scanner data. The main threats are (1) that the analysis hinges on precise measurement of prices, but prices in scanner data are the division of weekly revenue by weekly number of units. Thus each observation is a quantity-weighted weekly average of possibly different prices. The weekly average price is a mix of different prices if, for example, the price changed midweek, some consumers used coupons but not all, or club members faced a lower price; (2) that some price endings are also tied to specific promotion strategies, acting as non-price demand shifters that are hard to control for; (3) that price *endings* are chosen endogenously, making a simple comparison between, e.g., \$4.99 and \$5 impossible to do; and (4) that as with every attempted demand estimation we worry about simultaneity and price endogeneity.

I deal with measurement error in scanner data with a new data cleaning approach, and with endogeneity by using instrumental variables. I restrict attention to what are likely to be observations of "real" prices (prices consumers actually paid and not averages of different prices paid during the week) by excluding obvious averages where the price is not to the cent, and by using the length of same-price spells. I argue that

---

[2]This holds true even if consumers round up when they ignore the low digits.

selection on price-endings is not biasing the estimation but limits variation in the data. I use multiple fixed effects to control for most classic endogeneity worries, and supplement it with instrumental variables as in Hausman (1996). The price instruments are the concurrent average price of the same product in stores that share a distribution center but are located in different cities. This is a slight improvement over the classical Hausman Instruments in that "other" stores are selected based on a shared costs component (common distribution center).

Following these data cleaning procedure and estimation technique, the drops in demand are clear, as shown for thousands of products in Figure 3. That is, demand is flatter across cents within a dollar-digit but is dropping when the dollar digit changes. For example, a change in price from, say, $4.99 to $5.00, is associated with a drop of about 3-5 percentage points in the number of units sold.

These drops, together with the price level and price elasticity, can then be translated to approximate the bias parameter. The bias parameter is by how much the price should change absent the bias to generate the same change in demand as from the discontinuous drop. Namely, consider a drop of 4.5 percentage points in demand crossing from $4.99 to $5.00, and an estimated price elasticity of -1.5. Then, a change of 4.5% in demand can also be coming from a 3% change in price. 3% of $5 are 15 cents, implying a bias of 0.15, since for these parameters the drop in demand is the same as increasing the price from $4.85 ($\approx (1 - 0.15) * 4.9\dot{9} + 0.15 * 4$) to $5.00 absent a bias. I then estimate for each product the drops in every dollar digit, and its own price elasticity (using instrumental variables), to calculate the bias. While this method provides an insight about what identifies the parameter, an alternatvie is to estimate the model in one step with non-linear least squares. In essence, I regress log-quantity on log-*perceived*-price, $log\left((1 - \theta)\,p + \theta\,\lfloor p \rfloor\right)$, and other controls and fixed effects. As a third alternative and robustness test, I estimate a log-linear model, that also generates measures of the same bias parameter but assumes a different demand structure. The different estimation techniques result in similar estimated parameters, with rank correlation of 0.91 between the NLS and log-linear estimates, and 0.7 between the drops-based estimates and log-linear estimates. The average bias level is 0.16-0.24 depending on the estimation technique. Meaning, consumers ignore on average about 20 cents of a 99-ending price. I also estimate a specification with homogeneous left-digit bias and semi-elasticity, and estimate a bias parameter of $\theta = 0.16$ (0.01).

I then take the product-by-product estimates of left-digit bias and ask what product and clientele characteristics are correlated with higher bias. If the bias is driven by relative thinking (Tversky and Kahneman (1981)), it should be increasing with price since 10 cents out of $2 might feel relatively more important than 10 cents out of $8. This form of thinking is irrational if there is one unit purchased from each product. However, such a finding might be misleading if more expensive items are also purchased less often than cheaper products. That is, it can be a manifestation of rational inattention if the bias is lower for more frequently purchased products, since each per-unit cent weighs more heavily in the overall budget. Taking the effects together, a 10% higher priced item is associated with a 0.02 higher bias (meaning, ignoring additional 2 cents out of a 99 ending price), and an item purchased with 10% more units per trip is correlated with 0.02 lower bias.

In the first part of the paper I establish the bias existence, and estimate its magnitude. In turn, it allows to ask how should and how do firms price in response. Given that I find large discrepancies between predicted and actual pricing patterns, I ask what are the implications of these discrepancies.

The model predicts an optimal price given parameters of left-digit bias, elasticity, and cost. Discontinuities in demand lead to a pricing schedule by monopolistic firms that has distinct features (also for multiple-products monopolist in Appendix C, or price-competing firms, in Appendix B). Given a continuous cost distribution, prices will bunch just below round prices (e.g, bunch at $4.99), and this excess mass will be

coming mostly from *above*. That is, there will be a well-defined region of missing prices above the 99-ending price (but not below). While the prevalence of 99-ending prices has captured the attention of many researchers, the latter prediction has not yet been explored, and separates this model from the few other models suggested in the literature.

Qualitatively, firms pricing behavior follows the predictions and comparative statics of the model. Not only do retail prices end with 99 (24%-34% of popular products) but they rarely end with a price ending lower than 19. When they do, they are almost always on-sale prices, and likely an artifact of specific promotion strategies (such as "2 for $6"). A third prediction of the model is that the range of missing prices is increasing with the dollar digits. For example, under the model, the next-lowest price above $3.99, with left-digit bias of 0.05 and elasticity of -2, should be $4.65 while the next-lowest price above $4.99 should be $5.69. This is supported by firm pricing behavior, as low price endings are used less and less as the dollar digit increases.

However, quantitatively, there is a strong discrepancy between what the estimated bias imply for firms pricing patterns and what they actually do. Under the model, the estimated levels of left-digit bias (0.25 or even 0.16) and price elasticity imply that *all* prices should end with 99. Either more elastic demand or lower bias decrease the next-lowest prices and hence lower perceived bias or higher perceived elasticity can reconcile the patterns in the data. However, calibration exercises show that even for unrealistically high elasticities (such as -15), there are still big discrepancies. Since price elasticities are one of the key parameters the seller cares about in pricing, it is unlikely that they will perceive it to be so different. In converse, estimating the bias level is less likely and not straightforward to do even if the firm is aware of it.

I estimate the firm's *perceived* left-digit bias and elasticity that rationalize their pricing behavior under the model. The estimation technique allows to estimate the perceived price elasticity solely from the price distribution, even absent knowledge of quantities or costs. The estimation technique relies on strong assumptions, and the estimates are noisy, but reasonable. Compared to demand-side estimates, I find perceived elasticity to be more elastic at -3.35, and the perceived left-digit bias to be much lower at 0.023. Depending on assumptions, I can reject a perceived bias larger than 0.043, which is lower than the demand-side estimated levels by a few factors.

Finally, given the demand side estimates, I ask what are the firm profits at different scenarios. Namely, I compare counterfactual profits from ignoring the bias or underestimating it to a benchmark of full optimization. Compared to ignoring the bias, profits for optimizing firms are higher by 3-5 percents. The result holds under various scenarios allowing for a multi-product price setting, or for heterogeneity in left-digit bias, perceived bias, or price elasticity. For example, for the main point estimates, gross profits are higher by about 4%. Meaning, that at the very least, the mere existence of the bias can have large effects on profits. But perhaps, the firms respond enough as to recoup most of these profits. The most interesting case is when the firm prices *as if* the bias is much weaker, as it seems to do. Still it loses more than half - about 1-3 percents - of *gross* profits. The loss occurs regardless of whether consumers adjust prices downward or upward when they ignore the lower digits.

How can firms respond to the bias but at the same time underestimate it? First, this is not the first paper to document firms making consistent deviations from profit maximization (e.g. Bloom and Van Reenen (2007); Goldfarb and Xiao (2011)). Yet, a specific explanation for the partial correction, is that a partial response to the bias, such as changing $2.00 to $1.99, is both more "intuitive" and brings large gains, while a full correction, such as changing $2.40 to $1.99, is counter-intuitive and less impactful on profits (and hence also harder to pin down with noisy data). In the final section I discuss alternative explanations, such as the model being wrong or the plausible case in which consumers' bias level responds to the pricing behavior of the firm. I argue that these can not be ruled out, but seem unlikely.

This paper adds to the literature on 9-ending prices, on left-digit bias in general, to the literature in behavioral industrial organization, and on retail demand estimation.

Of course, the prevalence of 9 ending prices is not new, and has been extensively explored. Some papers draw attention to documenting the fact that many prices end with 9 (Hackl et al. (2014); Levy et al. (2011)) or revert to 9 following currency changes (Aalto-Setala (2005); El Sehity et al. (2005); Shlain (2018)), as a "revealed preference" argument for its benefits – firms use 9-ending prices because it is beneficial for them. Many papers test the effect of 9-ending prices on demand in experimental data in the lab and explore underlying mechanisms (Bizer and Schindler (2005); Carver and Padgett (2012); Schindler and Wiman (1989); Schindler and Kirby (1997); Schindler and Chandrashekaran (2004); Thomas and Morwitz (2005))[3]. Fewer papers examine the effects of 9-ending prices on demand, either in field experiments (Anderson and Simester (2003); Ashton (2014); Bray and Harris (2006); Dalrymple and Haines Jr (1970); Dube et al. (2017); Ginzberg (1936); Schindler and Kibarian (1996))[4], or in observational data (Blattberg and Wisniewski (1988); Hackl et al. (2014); Kalyanam and Shively (1998); Macé (2012); Stiving and Winer (1997))[5]. These papers find that 9-ending prices increase demand, but the estimates are imprecise. This paper adds to this literature by providing visual reduced-form evidence for the bias using big data, and a structural estimation and interpretation of the findings allowing to import the bias into the firm's problem[6]. This is also the first paper that quantifies firm response to the bias.

On the formal theory side, few models relate directly to 9-ending prices (Basu (1997, 2006); Stiving (2000)), and some bounded-rationality models generate discrete price-setting behavior (Chen et al. (2010); Gabaix (2014); Matějka (2015)). This paper provides a portable model that is straightforward to take to the data, connecting both the consumer and firm sides. A predecessor smooth model of left-digit bias (DellaVigna (2009), the one used here on the demand side) was estimated in various economic settings, but not on prices (Lacetera et al. (2012); Busse et al. (2013)). This difference is crucial, since prices are manipulateable[7]. This paper estimates left-digit bias in an everyday environment of grocery shopping, and adds a way to do so in face of strategic response of firms.

The paper relates to the literature in behavioral IO in two senses. First, by studying a market where firms respond to biased consumers (e.g., Akerlof and Shiller (2015); DellaVigna and Malmendier (2004, 2006); Heidhues and Kőszegi (forthcoming); Spiegler (2011)), and second by studying a case where the firms are also "behavioral", in the sense that they are not profit-maximizing in a systematic way (Cho and Rust (2008, 2010); DellaVigna and Gentzkow (2017); Goldfarb and Xiao (2011); Hanna et al. (2014)). A companion paper, Shlain (2018), exploits a natural experiment in Israel and finds that firms do respond to left-digit bias in the long-run but in a way that is consistent with relearning a rule-of-thumb behavior rather than understanding the underlying model.

---

[3]Common techniques are to do "recall" tests, and find that subjects are worse at recalling a price ending with non-round numbers. A summary table is available in Carver and Padgett (2012), and the most recent evidence, by Snir et al. (2017), is also exploring consumers response to the same policy reform I study in a companion paper (Shlain (2018)). Other explanations refer to drop-off mechanisms, such as in Bizer and Schindler (2005) who find that subjects purchase more out of the same budget when prices are odd.

[4]Anderson and Simester (2003) and the earlier papers in that list run experiments with varying prices in mail-in catalogs. An exception of supermarket setting is Ashton (2014) who uses the data from Chetty et al. (2009) and shows in reduced form that all of the response to experimentally-included sales tax is coming from prices who after the inclusion change their dollar digit versus those that did not.

[5]Observational data papers usually find higher demand for 99 or 9-ending prices by including a dummy for these. Surprisingly, these papers rarely discuss the issue of price averaging that is common across these data. The paper by Stiving and Winer (1997) is doing a beautiful work in testing various hypotheses of the mechanisms supporting 9-ending prices, but is limited to only 4 products with low prices (under $1).

[6]A few papers document that prices tend to end with 9, or 99, and ask questions on the effect of this phenomenon on price stickiness (Anderson et al. (2015); Ater and Gerlitz (2017); Levy et al. (2011)) and the pass-through of taxes (Conlon and Rao (2016)). This paper adds to this literature by micro-founding this stickiness as an optimal pricing response of firms.

[7]Indeed, two papers that examine left-digit bias in house prices, find extensive sorting of prices, but try to argue that the price ending is chosen as-good-as random, which is clearly not the case in retail pricing (Chava and Yao (2017); Repetto and Solis (2018)).

Finally, the paper relates to the vast literature of product demand estimation (see a review by Nevo (2011)), by contributing a novel method to estimate demand solely from the price distribution. The method exploits the discontinuities created by the bias, importing insights from public economics (Saez (2010)). Thus, this paper also contributes to the recent literature in structural behavioral economics with a rare case in which a behavioral bias allows identification rather than make it impossible (see review by DellaVigna (forthcoming)).

To conclude, in this paper I document a bias that is prevalent in everyday decision making of consumers. Even though the market reponds to the bias and exploits it, it is not sufficient to drive it out of existence. More suprisingly perhaps, firms do not seem to fully exploit the bias, leading to significant losses on their side.

The rest of the paper is organized as follows: Section 2 describes the model with left-digit biased consumers and solves the problem of an optimally pricing monopolist. Section 3 describes the data and cleaning procedures. Section 4 provides reduced form and structural estimates of the demand side in light of the model. Section 5 then moves to the firm side, analyzing firm pricing facing the estimated demand. Section 6 provides counterfactual exercises regarding firm profits. Section 7 concludes with a discussion of the plausibility of the results.

2   Model

Existing Theories   Most of the literature on 99 pricing is not formally modeled. The literature assumes that firms price at 99 because they face some non-standard consumers' response to pricing. Main explanations are in the form of "level effects", where consumers drop-off the lowest digits of a price, and "image effects" where the price carries a signal about the product, either about its quality or price (see Stiving and Winer (1997) for a review). The model that will be described shortly is a form of "level effect", but it also has something to say about image effects.

A few formal models do exist. Most notably, Basu (2006) models rationally inattentive consumers choosing whether to observe the lower digits or ignore them altogether, and then replacing them with the mean price ending (conditional on the first digits). Basu finds multiple equilibria that are either complete ignorance (with firms responding by solely pricing with 99), or full attention. Other papers, assume bounded rationality on the side of consumers or firms, which in turn creates discrete pricing, but not necessarily at 99 (Chen et al. (2010); Gabaix (2014); Matějka (2015)).

Another popular class of models is of inattention to certain components of a price (such as sale taxes in Chetty et al. (2009) or shipping fees in Hossain and Morgan (2006)). The model was later translated to inattention to lower-digits by Lacetera, Pope and Sydnor (2012) and estimated as the effect of odometer readings on prices in the used cars market. This model of consumer behavior, which I use and complement with a responding pricing firm, posits that each consumer reads a number left to right, and has a probability θ of ignoring the rightmost digits. In Lacetera et al. (2012), that number is the odometer of the car, and in this paper it is the price of a product.

Model   The model consists of left-digit-biased consumers and monopolistic firms[8]. Consumers choose a quantity to purchase to maximize utility, but may misperceive the price. Firms price to maximize profits,

---

[8]In appendix B I solve a case of firms pricing under price-competition with left-digit biased consumers. The case of multi-product monopolist is in appendix C.

according to their beliefs of consumer misperception. Misperception is a primitive of consumer behavior, and is not affected by firm behavior. In the last section of the paper I will discuss how crucial this assumption might be.

Consider a product whose price is p. Assume that a consumer perceives the price as $\hat{p} = \hat{p}(p; \theta)$, a distortion function that causes the consumer to perceive the price as possibly different than it actually is. Given this paper's empirical context, the $\hat{p}$ function is such that consumers observe the left-most digits in full, but observe the cents digits with a lower level of attention. Specifically, a left-digit bias parameter $\theta \in [0, 1]$ has two interpretations[9]. One interpretation is that it describes that each consumer is putting a relative weight θ on the lower digits of the price, while a second alternative is that θ share of consumers ignore the lower digits altogether. Though different in interpretation, these two modes lead to similar qualitative predictions (see Appendix B for when they differ). The former interpretation is a bit simpler to solve algebraically, and also captures the observation that "4.99 feels lower than 5.00" at the person-specific level. This is the model of choice. That is, from now on assume that

$$\hat{p} = \hat{p}(p; \theta) = (1 - \theta) p + \theta \lfloor p \rfloor \tag{1}$$

A crucial modeling choice made here is that if lower digits are not observed, they are completely dropped. This is important for welfare and total profits (do consumers over- or under-consume?), but can not be identified in the data. In addition, it is less consequential for the impact of misoptimization by the firm. I highlight when this assumption is consequential and discuss alternatives.

If the agent has utility of the form $U(y, q) = y + \frac{\left(\frac{q}{A}\right)^{1+\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}}$ where y is the residual income, q is the quantity purchased of a good, and A is translating quantity to numeraire value, then overall demand will be of the form[10]

$$D(p; \theta) = A\hat{p}^{\epsilon} = A\left((1 - \theta) p + \theta \lfloor p \rfloor\right)^{\epsilon} \tag{2}$$

To get some intuition, the impact of the θ left-digit bias parameter can be understood by observing three distinct consequences for price-setting and demand: First, the effect of a 0.01 price increase on demand is lower than in the standard model if only one digit is changed (i.e. 2.95 to 2.96); Second, the penalty of a 0.01 price increase is higher if higher level digits also change (i.e. 2.99 to 3.00). Finally, almost all prices (except round ones where $p = \lfloor p \rfloor$) appear to be lower than they are - hence for every price p that is not a round price, demand is higher when θ is higher (again, by a non-verifiable assumption).

Consider a monopolist who faces a demand function $D(p; \theta)$. The firm earns p per unit, and pays a fixed unit cost c. Then, the monopolist gross profits are

$$\Pi(p; \theta) = D(p; \theta)(p - c)$$

i.e., demand is driven by the *perceived* price while per unit profits are governed by the *true* price. This is a common feature of models with inattention (e.g. Farhi and Gabaix (2015)), where there is a discrepancy between true utility and maximized utility on the consumer side, causing distortions.

---

[9]In other papers, θ is termed "inattention". However, other forces might lead to the same behavior, such as imperfect price recall, price categorization, or a tendency to choose round numbers for internal values or reference-prices. Therefore, I use the term left-digit bias which captures all of these, and remain agnostic about the behavioral mechanism.

[10]Conversely, if there is type mixture as described above, then following the same formulation will lead to demand of the form $D(p; \theta) = A\left[(1 - \theta) p^{\epsilon} + \theta \lfloor p \rfloor^{\epsilon}\right]$.

In section 5 I consider a specific form of misoptimization by the firm. The monopolist understands the model and knows the true structural elasticity, but might be wrong about the level of left-digit bias, so will maximize profits given their belief. i.e., it maximizes $\Pi = D(p; \hat{\theta})(p - c)$, where $\hat{\theta}$ might be different than θ. For now, assume that $\hat{\theta} = \theta$.

optimal pricing given θ   For ease of exposition, I decompose a price p into its decimal basis components, s.t. $p = p_1 + p_{0.1} + p_{0.01}$. For example, $p = 3.49 = 3 + 0.40 + 0.09$, so $p_1 = 3$.

With left-digit bias, we get a discontinuity in demand when changing p_1 since in the limit from below, demand is $A(p_1 - \theta)^\epsilon$ versus $A(p_1)^\epsilon$. To simplify the solution, assume that prices can be chosen from all real numbers between a natural number and a .99 ending number - $\underset{q_1}{\cup}[q_1, q] \subset \mathbb{R}_+$ where $q_1 \in \mathbb{N}$ and $q = q_1 + 0.9 + 0.09$. Proofs to the following propositions and corollaries appear in Appendix A.

Solving for optimal pricing yields the following:

Proposition 1. *Optimal pricing formula, for small θ.*

For any cost c and parameters ϵ and θ find the appropriate .99 ending number $q = (q_1, .9, .09)$ such that $c \in \left[\underline{c}_q, \underline{c}_{q+1}\right] \triangleq \left[q\left(1 + \frac{1}{\epsilon}\right) + \frac{\theta}{1-\theta}\frac{q_1}{\epsilon}, (q+1)\left(1 + \frac{1}{\epsilon}\right) + \frac{\theta}{1-\theta}\frac{q_1+1}{\epsilon}\right]$. Then, the optimal price for that cost c is

$$p(c; \theta, \epsilon) = \begin{cases} q & \text{if } c \in \left[\underline{c}_q, \bar{c}_q\right] \\ \left(c - \frac{\theta}{1-\theta} \cdot \frac{q_1+1}{\epsilon}\right)\frac{\epsilon}{1+\epsilon} & \text{if } c \in \left[\bar{c}_q, \underline{c}_{q+1}\right) \end{cases} \tag{3}$$

Where $\underline{c}_q$ and $\underline{c}_{q+1}$ are defined above, and $\bar{c}_q$ is defined below with an implicit equation as the minimal cost for which it is profitable to price strictly above q.

To make Equation 3 clearer I elaborate on the three threshold costs $\underline{c}_q$, $\underline{c}_{q+1}$, and $\bar{c}_q$. The former two, $\underline{c}_q$ and $\underline{c}_{q+1}$, are the costs for which the monopolist's profit is maximized as an *interior* solution at q and $q + 1$ respectively. i.e., where the first order condition is satisfied at q and $q + 1$. The image of costs in $\left[\underline{c}_q, \underline{c}_{q+1}\right]$ is hence prices in $[q, q + 1]$ which allows to use q and q_1 in the two cases in Equation 3. The third threshold, $\bar{c}_q$, is the cost for which prices will switch between q to a point in the segment $[q_1 + 1, q + 1]$. At $\bar{c}_q$ two conditions are met: (1) the profit is maximized (with an internal solution) on some price $P_q \in [q_1 + 1, q + 1] > q$; and (2) profits are equal at that price, P_q, and at q^{11}. For ease of notation, I omit q from P_q in the following:

$$\bar{c}_q = \begin{cases} P + \frac{(1-\theta)P + \theta P_1}{\epsilon(1-\theta)} \\ \frac{D_P P - D_q q}{D_P - D_q} \end{cases} \tag{4}$$

Definition 1. The *Next-Lowest Price P* is the lowest price used above a 99-ending price q, and is a function of the parameters θ and ϵ as defined by the following implicit equation:

$$P + \frac{(1-\theta)P + \theta P_1}{\epsilon(1-\theta)} - \frac{((1-\theta)P + \theta P_1)^\epsilon P - ((1-\theta)q + \theta q_1)^\epsilon q}{((1-\theta)P + \theta P_1)^\epsilon - ((1-\theta)q + \theta q_1)^\epsilon} = 0 \tag{5}$$

This equation is not analytically solvable, but can be easily solved numerically. Given P, we get $\bar{c}_q$ from either part of Equation 4. It means that no prices are set between $q_1 + 1$ and P.

---

[11] For very low θ, there is no gap, and $P_q = q_1 + 1$.

To gain some intuition regarding Equation 3 note that this pricing behavior is different than a no-bias world in two ways. First, the top case shows that prices bunch at 99, and second, the bottom case describes interior solutions as a modified markup rule with an added component driven by the bias.

The top equation shows that prices are set at 99 ending prices (q) for a range of unit costs, and hence with varying markups. This is an admittedly unsurprising, and a motivating prediction, of non-zero mass at 99-ending prices. The mass is caused by the discontinuous drops in demand. Thus, this model generates price stickiness that is not driven by frictions on the supplier side.

The lower expression in Equation 3, describes the interior solutions, with the lowest price being P_q. It shows that the optimal price is a modified Lerner equation with an extra term $\left(-\frac{\theta}{1-\theta} \cdot \frac{q_1+1}{1+\epsilon}\right)$. In the interior solutions region, the price is slightly higher than it would have been absent the bias (recall that $\epsilon < -1$). This is due to the left-digit bias making demand less elastic for changes within the same first digit. This added markup is increasing in the bias (θ), and the nominal price (q_1), and decreasing in the absolute elasticity (ϵ).

The pricing schedule is illustrated in Figure 1 showing price as a function of cost with left-digit bias. The diagonal gray line is the price without bias, and the thick lines are the optimal prices with the bias, exhibiting regions of costs priced at 99-ending prices and ranges of missing prices with low price endings.

Notice that in Figure 1 products whose price is 99-ending are usually (depending on the cost distribution) priced with *lower* markups. This can be seen as a "level effect" micro-foundation of the "image effect" style of explanations – 99-ending prices are indeed, in expectation, of higher value (insofar as costs are positively related to value).

Most important are the testable predictions of the model: That there is bunching at 99-ending prices, and that this bunching is drawn from missing prices above the round price thresholds.

Comparative statics We are interested in the effect of the elasticity and left-digit bias on the next-lowest-price P, and can show the following results:

Proposition 2. *Comparative statics of the Next-Lowest Price: For $\epsilon < -1$,*

The next-lowest price is lower when demand is more elastic, i.e. $\frac{\partial P}{\partial |\epsilon|} < 0$

For $q_1 > 0$, the next-lowest price is higher when there is more bias, i.e. $\frac{\partial P}{\partial \theta} > 0$.

The first part of the proposition shows that more elastic demand leads to lower next-lowest price. Higher elasticity means lower markups, so costs that go to a 99-ending price q are now higher. If the indifference cost $\bar{c}_q$ is kept fixed, and since $\bar{c}_q = P\left(1 + \frac{1}{\epsilon}\right) + ...$, it means that P must decrease. Of course, the cost is increasing, but the proof shows that it increases by less than offsetting this primary effect.

The second part of Proposition 2 states that higher bias leads to higher next-lowest prices. The effect of higher bias is simpler to understand. With higher bias, demand is rotated up, meaning it is higher everywhere in proportion to the price-ending (kept fixed at round prices). It makes the effect on higher-digit ending prices stronger than on low-digit ending prices. Higher demand at q increases the indifference cost and P, while higher demand above q decreases the indifference cost, lowering P. Because the effect is stronger on q than on lower-ending prices, the net effect is to increase P.

In the data we observe prices across time, products and stores. So far, we have not made any assumptions about what drives variation in prices – price variation can be generated from differences in costs, in

elasticities, or in bias levels. Like common in the literature (e.g. Hausman (1996)) I will assume that for a specific level of observations (e.g., same brand and chain, across stores and time) elasticity and bias are fixed, and the price-setting variation is driven by changes in costs[12]. Note, that apart from having a support that includes $\underline{c}_q, \bar{c}_q$ (i.e., the costs that generate both a 99-ending price and the next-lowest price above it), we do not need to make any assumptions about the cost distribution, $c \sim F$, to test the NLP predictions in Proposition 2.

A simple corollary is showing that the right-digits component of the next-lowest price is increasing from one threshold to the next:

Corollary 1. *For $\epsilon < -1$, the lower digits of the Next-Lowest Price (i.e., $P - P_1$) are increasing with the first digits.*

This corollary means that more nominally expensive products have larger regions of missing prices. For example, for the same parameters of left-digit bias and elasticity, the lowest ovserved price beginning with 4 should be \$4.30 and with 5 is \$5.32.

This is not necessarily tied to larger shares of 99-ending prices, because excess mass depends on the shape of the underlying cost distribution. The share of observations in a 99-ending price q is $F\left(\bar{c}_q\right) - F\left(\underline{c}_q\right)$. Higher left-digit bias or less elastic demand lead to greater $\bar{c}_q$, but both also shift the costs that go to 99, meaning also higher $\underline{c}_q$. It can be shown, even for the simplest uniform cost distribution that the sign of effects on share of 99-prices depends on the parameters. For example, if costs are uniformly distributed, then a change in the excess-mass due to an infinitesimal increase in θ is proportional to $\frac{\partial P}{\partial \theta} + \frac{1}{\epsilon}\left(\frac{\partial P}{\partial \theta} + \frac{1}{(1-\theta)^2}\right)$. If $\epsilon \to -1$ then this is negative, implying a decreasing mass at q. If $\epsilon \ll -1$ then the first term, which is positive as per Proposition 2, dominates, making the effect positive overall. The same kind of exercise holds also for the comparative static effects of the price elasticity[13].

In other words, unlike the next-lowest price predictions, the excess mass is sensitive to the shape of the cost distribution and the exact values of parameters. This means that point identification of left-digit bias and price elasticity from next-lowest prices and excess mass is theoretically possible, in the sense that the parameters have differing effects on the moments. This avenue is then explored and developed in Section 5.

3 Data

This project relies on various data sources, aiming to provide a thorough investigation of the bias. The ideal data include a large number of observations of precisely measured prices and quantities, where prices change due to supply shocks and all demand shifters can be controlled for. For validity purposes, we would also like the data to reflect many types of products, consumers, and retailers. As is often the case, ideal data are hard to find. Nielsen scanner data are great for validity, capturing a large share of products and retailers. It is also a long panel, helping with estimation precision. However, it has a few shortcomings mainly on the side of price measurement. Prices in the data often represent an average of different prices paid within a week, but not a price that any consumer actually paid. In contrast, data from a large US grocer have several advantages in measurement and separation between regular and on-sale prices, but represent only this retailer and their clientele.

---

[12]DellaVigna and Gentzkow (2017) show that there are different elasticities between stores within a chain-product, but that these differences are not translated to variation in prices; bias heterogeneity is also of course probable, but it seems natural to establish the mean bias as a first step (compare the mean-bias Chetty et al. (2009) to the heterogeneous-bias Taubinsky and Rees-Jones (2016)).

[13]For elasticity with uniform cost distribution the first derivative of excess mass at q to changes in ϵ is $\frac{\partial P}{\partial \epsilon}\left(1 + \frac{1}{\epsilon}\right) - \left(P - q + \frac{\theta}{1-\theta}\right)\frac{1}{\epsilon^2}$.

Therefore, the main dataset I use is the retailer, allowing for cleaner empirical analysis on a large set of products. I then use a sample from Nielsen as a validity and robustness test, and find similar results across more than 60 different retailers.

3.1   Retailer

The first dataset is scanner data of sales of a US national grocer, with a relatively rich set of variables. The same data were used in other papers such as Eichenbaum et al. (2011); Gopinath et al. (2011). Data are collected from the cash registries of the chain's stores and aggregated at the weekly level. Each observation is a distinct product (UPC) in a store in a week. As shown in column (1) of Table 1, there are about 74000 distinct products, in 250 stores, and 177 weeks (2004 to mid-2007). The data describe sales of more than 3 billion units, and $7.2 billion worth of sales.

Available variables for each observation include the number of units sold, the net revenue (the actual amount paid) and the gross revenue (which is the amount the store would have collected if each unit were priced at its regular price, absent discounts), among other variables (such as two cost measures - wholesale prices and adjusted gross profits). The price is then the division of net revenue by units sold, and the regular price is the division of gross revenue by units sold.

The data are used to conduct demand estimation, and therefore we might be worried about price endogeneity. Ideally we should get prices to vary due to cost changes (supply shocks), but not due to changes in demand. I take a common approach to address this issue, with instrumental variables for prices. I am using leave-out-prices (in the spirit of Hausman et al. (1994); Hausman (1996); Nevo (2001)) for the same product at the same period in other stores. To increase the likelihood that these other prices capture supply shocks rather than local demand shocks, I use stores in other cities, that share the same distribution center. This is another advantage of the data, where each of the 250 stores is assigned to one of 17 distribution centers supplying it the goods. I discuss the identifying assumptions with this approach when I describe the estimation in Section 4.

An important issue with such scanner data is that if a product did not sell in a certain week, its price is not observed (because zero units sold for zero revenues). It is likely that observations are not missing at random. That is, if a product is less likely to sell when its price is high, then there is a potential for underestimation of price elasticity. To mitigate this issue, I select popular products. I select all products that were available at regular prices for at least 50% of all observations (among weeks and stores in which they were at any point sold) and with at least 10,000 observations. To get some level of minimal price-ending and dollar variation, I am selecting the 1710 products whose regular prices end with 99 at most 75% of observations, or are all within a single dollar digit at most 75% of observations, as shown in column (2) of Table 1.

As shown in column 2 of Table 1, this procedure leaves me with 1,710 products and 51.3 Million observations. Since these are of the most available and popular products, while they are 2.3% of all possible products, they capture 19.2% of all revenue in the data. The average price is $3.42 (with standard deviation of 1.63), and the inter-quartile range of prices is $2.29 to $4.19. Meaning, most of the prices have exactly one dollar digit, limiting exploration of the effects of more or fewer digits (96.35% of observations are strictly between $1 and $10).

The reader might worry that I am selecting the sample based on outcome-related variables. Hence, I also present results for the complementary set of popular products, namely those with highly concentrated

prices, as a robustness[14]. Another limitation of the data is that it is not possible to infer the unit price for products paid by weight, such as fresh produce, meat, or deli items. I exclude these product categories[15].

In order to investigate the relationship of demand to changes within, versus across, dollar digits, the exact price paid is crucial. However, the inferred price (net revenue divided by units sold) does not necessarily, and often do not, describe the actual price paid in store. This happens when some consumers pay with coupons, if some have club membership specific discounts, or if some products are priced with non-linear pricing (such as "buy 1 item get the second for 50%"). Another potential source of multiple prices, which is when prices change during the measurement week, does not occur in this data set, since the measurement weeks are aligned with the price changing frequency. However, in AC Nielsen RMS data this is a sometimes severe issue as will be described next. The nature of averaged prices in scanner data is a known issue in the literature, and some of its implications on inference regarding economic phenomena, such as demand estimation and price changes, were explored (e.g. Einav et al. (2010); Eichenbaum et al. (2014)). The retailer data have a unique advantage since there is no rounding. e.g., a price in the data, which is the result of dividing revenue by units, might be $3.044. In such cases, when the price is not "to the cent" then it must be the result of weighted averaging of several prices (here, for example one possible mix can be 41 units selling for 2.99 and 9 units for 3.29). Therefore, I exclude these observations, reducing the number of observations by about 10%, to 46 Million (column (3) of Table 1). However, this might keep other observations that are a mixture but happen to have an average that is to-the cent. For example, 9 units purchased at $4.99 and 6 at $4.49 will give an average price of $4.79[16]. A different approach is to use the length of the price spell. If the same price appears for two consecutive weeks or more, it is less likely that this price is non-real. Indeed, as shown in Figure A.3 by the dark circles, already at two-week spells, most prices are to-the-cent. This method is mostly useful for Nielsen data, where all prices are rounded to the cent, so the former approach is unavailable. Finally, a sure way to keep only non-averaged prices is to only consider the instances where the net price equals the regular price (which happens for about two thirds of observations). Regular prices do not change very often, and are indeed always to-the-cent.

Some statistics about the price endings are shown in Panel B of Table 1. When I include prices that are not to-the-cent I round the prices, emulating what happens in the Nielsen data. As we move from all observations to the "real prices", the share of 9-ending prices increases, meaning that other price endings are more likely to be the result of averaging of multiple prices. Notice also the higher share of 0-ending prices for all prices (column (3)) versus only regular prices (column (4)). Meaning, 0-ending prices are strongly positively correlated with being on-sale[17].

If 9-ending prices are mostly not on-sale, while 0-ending prices (and others) are on-sale, it will lead to spurious higher demand for 0-ending prices without flexible enough controls of sales effects or if the model is misspecified. Prices are only one determinant of demand, and different promotion techniques can lead to

---

[14]The other 2149 products, are characterized by having just as high availability and 10000 observations, but have either more than 75% of observations ending with 99 (only 125 products), or more than 75% of observations starting with the same dollar digit. This mostly happens for cheaper products, priced under $1. There are 55M observations, representing 170M units sold annually (more than the main sample) with annual revenue of $292M (less than the main sample). Indeed, the average price is $2.2 (versus $3.5 for the main sample), with fewer 99 and 9-ending prices (15.4% and 63.7%).

[15]I exclude the following: Delicatessen, Food Service, Fresh Produce, Meat, Seafood, and Alcoholic Beverages. This is a serious issue since weight adjustments only work for some products but not for others. For example, sliced turkey breast, packaged in the deli and priced by weight with a sticker on the wrap will appear as a single unit.

[16]Under the assumption that there are at most two prices paid, with one of the prices being the regular price, the other price and quantity purchased at each price can be inferred in theory from the average price, the total number of units, and the facts that prices should end with a cent. However, this is imperfect and allows for multiple prices, and requires some assumptions to restrict possible prices to a single price. I therefore prefer to err on the safe side by reducing sample size.

[17]Note that these findings stand in contrast with the assumptions of "image effects". First, under image-effects rationalization, round prices should be linked to high-quality items, but at the product level, they occur mostly when the product is on-sale. Second, a common explanation for 99 prices effectiveness in this literature is that they are learned through consumer experience to signal a discount. But, the contrary is true, as 9-ending prices, if anything, should signal the regular, higher, prices.

stronger effects than others. Indeed, as casual observations suggest, round prices are commonly the result of quantity sales (e.g., "2 for $5" which means an individual unit price of $2.50) that might have stronger effects than equally sized sales (e.g. Blattberg and Neslin (1990); Wansink et al. (1998)). Appendix Figure A.2 shows how the shares of observations with to-the-cent prices (left panel) and on-sale prices (right panel) depend on the last digit of the price. Indeed, almost all 9-ending prices represent prices actually paid in the store (98%), but only 9.9% (!) of them are on-sale. Compare that to 0-ending prices, that are also mostly real (92%), but 91% of them are on-sale prices.

3.2 Nielsen

Nielsen Retail Scanner (RMS) and Consumer Panel (HMS) data are provided by the Kilts Center at the University of Chicago. Like the retailer data, RMS records weekly UPC-store level quantity and revenues. Overall the scanner data record observations for over 35,000 stores during 2006-2014 in the US, regarding about a million different unique products. As per this version, I am using the Nielsen data as a robustness test using a sample of products and stores selected by DellaVigna and Gentzkow (2017) (see for details on data selection), of 12 items (termed "high-end" in their paper) in food stores. The summary statistics are described in Appendix Table 1. The sample consists of 28 million observations of these 12 products, in 11000 stores, with a total of 188 million units sold for $881 million. The data have a few disadvantages relative to the retailer data, which I elaborate on with possible solutions in Appendix Section D. In a nutshell, an effective way to identify sales and non-real prices in the data is by using the spell length of a price. If a popular product's price changes from week to week, it is likely to be a non-real price. If a product's price is kept constant for many weeks, e.g. for six consecutive weeks, it is likely a regular price. I discuss this idea further in Appendix D.

4   Demand Estimation

In this section I provide reduced-form support for left-digit bias, and structurally estimate the bias parameter as prescribed by the model. Finally, given estimates at the product level, I investigate how do estimates change with respect to characteristics of the product and its clientele.

4.1   Reduced-Form Evidence

Figure 2(a) illustrates constant elasticity demand curves with and without the bias, i.e. $\log Q = A + \epsilon \log\left((1-\theta)\,p + \theta\lfloor p \rfloor\right)$, for $\theta = 0.15$ on the left panels and $\theta = 0$ on the right. On the right, without bias, demand is a line with slope ϵ. On the left, the slope is flatter within each dollar digit, and demand drops when the dollar digit changes (highlighted by different colors). The dashed line is the fitted line from a log(quantity) on log(price) regression[18]. In turn, Figure 2(b) shows the residuals by price from these regressions. The sawtooth pattern of residuals for the bias case, with discrete drops at dollar thresholds, is a manifestation of the bias. If there is no bias, as in the right panel of figure 2(b), residuals are a flat line at 0. If the model is misspecified, it will exhibit some other general shape.

---

[18]The simulated price distribution is uniform. If it were the case that prices are more likely to be 99 ending, a regression line will be closer to the 99 ending prices. The slope of the regression line, without taking the bias into account, is interpreted as the price elasticity. The slope is affected by the distribution of price endings, but can go either way. For example, if prices are $4.99 and $5.05, the regression slope will be steeper, making demand seem more elastic than it is, while if prices were $5.05 and $5.99, the slope will be flatter and demand less elastic. I assume, as in the model, that the price elasticity is the demand response to true price changes.

In the following I focus on residualized demand by exact price. The idea is to net out any factors explaining differences in demand across observations that are not due to price-endings, such as seasonality, store and product characteristics, promotion mix, and price elasticities. This also allows to aggregate over more flexible demand structures, for example different elasticities and price distributions for different products.

Figure 3 shows that the sawtooth patterns are found in the data. Residualized demand is increasing within a dollar-digit and drops across digits. But before turning to the results, I would like to make the empirical strategy clear.

There are four main empirical threats to overcome. The unique challenge in this paper is that the analysis hinges on exact prices. Because prices are revenue over quantity, there is risk of prices in the data not representing the price paid. As mentioned above I mitigate that worry by excluding prices that are not to-the-cent and by excluding one-week spells, or at the extreme by keeping only regular prices.

A second issue is the need to control for the promotion mix. Normally, when a retailer is putting a product on sale demand will increase by more than the mere price effect. The retailer uses various strategies - different tags and postage, advertisement in other parts of the store or outside of it might be in place, and the promotion might be contingent on (or nudging for) purchasing multiple units. As shown in figure A.2, since 9-ending prices are usually regular prices and 0-ending prices are usually on-sale, proper controls for sale effects are required. I am using multiple fixed effects for that purpose - a dummy for whether the product is on-sale, 4 dummies for the last digit of the price (0, 5, 9, or others), and 5 dummies for the spell length (since sales are usually shorter price spells as shown in Figure A.3) – all proxy different aspects of sale strategies, but all missing important aspects. A more extreme solution is to conduct the analysis only on regular prices. This has the downside of capturing a different price elasticity, and cutting down more heavily on the data. I take both approaches. The main analysis is conducted with all prices and sales controls, but the results are indeed cleaner for the regular prices.

The third challenge is the sorting of *price-endings*. As long as prices change due to costs, and as long as there is enough variation in price-endings (guaranteed by sample selection) this is not a worry. The reason is, that I am estimating the demand curve, and while prices are a function of the costs, demand is not. From the consumer perspective, it does not matter that different costs translate to the same price. Meaning, the shape of the demand, by price, is not affected by sorting on price. Of course, if all prices end with 99 or all prices begin with the same dollar digit, there is no variation and no hope for separating the price-ending effects from price effects.

Since the data lack exogenous variation in prices, we worry of elasticity bias driven by unobservables (such as local demand shocks) affecting prices and quantity sold. However, much of the demand variation is absorbed by fixed effects, capturing seasonality, products characteristics, permanent store differences, and sale strategies. As Rossi (2014) argues, long panels with multiple fixed effects might not be subject to substantial bias[19]. Relatedly, DellaVigna and Gentzkow (2017) show that chains set uniform prices across stores, hence implying that price variation is not driven by local shocks and does not equates demand and supply at the store-week level. Meaning, the worry should not be too big. So, even though I use instrumental variables for the main results, in some cases I estimate a simple OLS (or a non-instrumented NLLS). I construct instrumental variables and find slightly more elastic demand, but small impact, if any, on the patterns and estimated bias. I instrument for $logP_{ist}$ with $logP_{id(s)t}$, the average log-price of the same item i in same period t in stores (s) in other cities but the same distribution center $d(s)$. The assumption here is that this average

___

[19]In fact, Hausman (1996) argues the following: "to the extent that supermarkets set their prices... under an assumption of constant marginal costs (in the short run) and do not alter their prices to equilibriate supply and demand in a given week, prices may be considered predetermined... then IV methods would not necessarily be needed".

price is capturing changes in price that are driven by cost changes. The costs are likely to be common across stores and hence are unlikely to be driven by local demand shocks[20]. These instruments can not fully resolve the issue of endogeneity, since it might be that there are aggregate demand shocks that affect prices across stores in a certain week.

Finally, cross-elasticities with other products are an important component affecting demand[21]. I am therefore including a variable of the average price of other products in the same category (as defined by the retailer, e.g., "tea"). The common approach in the literature is to add all "close" products flexibly (as in the lowest level of almost ideal demand systems, as in Hausman et al. (1994), or to use BLP-like random coefficients (as in Nevo (2001)), or to take a Bayesian approach (Rossi et al. (2012)). Since I am using thousands of products and tens of millions of observations, and since I am excluding some observations with non-real prices, these approaches are harder to implement in practice. By adding a single variable of mean log price, I essentially estimate the average cross-elasticity to products within the category (I find it to be on the order of 0.1-0.3).

I am using the following specification *product-by-product*:

$$logQ_{ist} = \epsilon_i logP_{ist} + \alpha_i logP_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu_{i,on-sale(ist)} + \mu'_{i,spell.length(ist)} + \nu_{i,cents.digit(ist)} + e_{ist}$$
(6)

Where i is product, s is store, and t is period. Q is number of units sold, and P is the price. ϵ captures the product-chain elasticity as the coefficient on log-price. I instrument for $logP_{ist}$ with $logP_{id(s)t}$, the average log-price of the product in stores in other cities that share the same distribution center. The second term captures cross-elasticities, where $logP_{c(i)st}$ is the average log-price of other products of the same product category $c(i)$ in the same store s at the same period t. β is a store fixed effect capturing a steady state product average demand in a store; γ is a year fixed effect capturing long-term shifts in demand; δ is month-of-year fixed effect capturing product specific seasonality; μ is an indicator for whether a product's price is on-sale or not (by comparing the actual price to the regular price, and calling it "on sale" if the actual is lower than the regular). μ' is a spell-length fixed effect for spells of length 2, 3, 4, 5 weeks, and more than 6 weeks, capturing on-sale propensity and some stockpiling effects; and ν is a fixed effect for the cents digit being 9, 0, 5, or other, capturing different promotional strategies that are correlated with specific price endings.

I then take the residuals, $\hat{e}_{ist}$, and regress them on price fixed effects at 10 cents bins (e.g., \$2.90-\$2.99 is one bin).

$$\hat{e}_{ist} = \sum_{\tilde{p}=\{1.0,1.1,...,11.9\}} \alpha_{\tilde{p}} \mathbf{1} \left(p_{ist,1} + p_{ist,0.1} = \tilde{p} \right) + u_{ist}$$

To get confidence intervals I bootstrap the two steps together 300 times, clustering by store. I do not do this in a single step due to the large number of fixed effects (about 470,000 fixed effects plus 3,420 coefficients), making it computationally inefficient.

The main results, for 1,710 products are shown in Figure 3. The horizontal axis is the price, and the vertical axis is the residualized demand. Each point is the estimated $\alpha_{\tilde{p}}$, the residual demand at a specific price (in 10 cent bins). The top panel shows the residuals from specification 6 on the main sample (column 3 in Table 1). The bottom panel shows results from a restricted sample of regular prices, and the patterns are cleaner there (column 4 in Table 1). The figures show that quantity purchased, netted of price, promotion, and seasonality

---

[20]The first stage for all products is very strong, with the mean coefficient across products being 0.86 (standard deviation of 0.14) with the smallest F-stat being 69 and most are on order of tens thousands. This strong relation is likely driven by zone pricing behavior, and also explains why a simple OLS provides almost identical results, just less noisy.

[21]Cross elasticities are also important for pricing. To keep the model solvable analytically, and allow for comparative statics, I ignored cross-elasticities on the theory and estimation of the firm behavior. I add it back in conducting counterfactual exercises.

effects, is increasing within each dollar-digit and drops at dollar crossings. Such patterns are consistent with demand of left-digit biased consumers. Meaning, after controlling for all other effects, higher price-endings are associated with higher demand. Under the model, this is driven by ignoring a fixed proportion of the price-ending, creating larger discounts of the total price for higher price-endings.

In the main sample, for some prices there is actually no drop crossing the dollar threshold but for the next 10-cent prices (for example, demand only drops at $5.10 rather than at $5.00). A simple comparison of demand at $4.99 to $5.00 would lead to the conclusion that there is opposite bias, suggesting that perhaps round prices have some innate appeal. However, as can be seen in the bottom panel focusing on the sample of regular prices, these patterns are driven solely by promotions. For the restricted sample of regular prices, the round prices are aligned with the overall sawtooth pattern and all drops are at round prices. It does mean, however, that despite my efforts, controls for on-sale effects on demand are not fully absorbed.

An alternative explanation for the sawtooth pattern is that firms price at lower-ending prices when there is a negative demand shock, and at high-ending prices when there is a positive demand shock (beyond seasonality effects). Yet, since the patterns are weakened due to sale strategies, and on the other hand are consistent for regular prices which are set for longer periods of time, this explanation seems less likely.

One interesting pattern is that the drop size does not decrease as the dollar digits increase. Decrease in drops is implied if elasticity and bias are fixed across products and prices. Conversely, fixed drops imply that the ratio of bias to elasticity is increasing. I will further explore these issues in the next subsections.

The patterns also go against the image-effects explanations, where 9-ending prices have some innate appeal. These models describe demand as a smooth function that is decreasing in price but exhibits discontinuous jumps of higher demand for 9- and 99-ending prices (e.g., "blip" description in Stiving (2000) and some interpretations of findings in Anderson and Simester (2003)). However, such behavior should appear in the data as flat residuals, with higher estimates only for the .9x fixed effects.

Next, I conduct several robustness tests to verify that the findings are not driven by the retailer, the demand structure, or the sample selection.

To verify that the behavior is not driven by the specific retailer, Appendix Figure A.4 plots the results of the same exercise for the Nielsen data, of 12 products from DellaVigna and Gentzkow (2017) sold by more than 60 retailers. I run the regressions at the product-retailer level. In Nielsen data, cleaning is based on spell-length only, and so are the on-sale proxy variables. As for the single item and retailer, the patterns are clear but noisier for prices with fewer observations (as can be inferred from the confidence intervals, common prices for these 12 products are at $3-$5 and $8-$10).

In appendix figure A.5, I explore the role of the demand structure. The patterns are not driven by the constant elasticity assumption, since the figure shows that a semi-elasticity specification produces similar sawtooth patterns (semi-elasticity implies elasticity that is linear in price). To estimate the patterns, I use the same specification and procedure as in Equation 6, replacing log-price with actual price (also replacing average log-price of other products with average price, and instrumenting with average price in other stores).

Appendix Figure A.6 provides robustness tests on a different sample, to show that these patterns are not driven by product selection. The figure shows the residuals from instances with regular prices for all remaining 2035 popular products (defined by at least 50% availability at regular prices and at least 10000 observations), with at least 75% of prices ending with 99 (merely 125 products) or at least 75% of prices starting with the same dollar digits. Meaning, popular products, but those with more concentrated price distribution, still produce the same sawtooth patterns.

In summary, patterns are very similar and consistent with the model across various specifications, products, and retailers. These patterns support the existence and soundness of the model. More is needed in order to translate these patterns to parametric left-bias estimates.

4.2 Structural Estimation of θ

I now turn to connect the reduced-form evidence with the model, and to estimate the left-digit bias parameter θ. This is a parameter of interest for multiple reasons. First, it has a clear interpretation in the sense of quantifying what fraction of the cents component of a price consumers ignore. Second, structural estimation allows to describe the demand curve with a functional form, and to find (analytical) optimal response for the firm pricing in response to it. And finally, structural estimation enables counterfactual exercises regarding firm profits under various scenarios.

As a benchmark, I estimate a homogeneous left-digit bias parameter with a semi-elasticity specification (details in section 4.2.3). I find a precise estimate of $\theta = 0.1635$ (0.0096) which means that on average consumers ignore 16 cents out of a 99-ending price[22].

For the main analysis, similar to the reduced-form results, I estimate the left-digit bias parameter at the product level. That is, I estimate multiple θ_is. I do so both for computational reasons[23], and in order to explore heterogeneity in the bias estimates. Some of the heterogeneity is coming from sampling bias. Even with tens of thousands of observations per product, there is only a small finite number of prices used, while for identification, multiple prices within and across dollar digits are required. Further, the specification is a simplification of the demand structure, especially lacking in controlling for promotion effects. But, this heterogeneity is strongly correlated with underlying characteristics of the different products and their clientele. Learning about this heterogeneity structure informs us about the underlying mechanisms driving the bias in the field.

4.2.1 Sufficient Statistics Approach

The bias can be inferred from the demand drops at dollar thresholds. For example, assume that demand drops by 4% when the price crosses the $5 digit. If elasticity is -2, a 4% change in demand can also be driven by a 2% change in price. 2% out of $5 is $0.10, so this implies a bias of 0.1, since the demand response in going from $4.99 to $5.00 is like a change from $4.90 to $5.00 if there is no bias. That is, as if the perceived price just below $5.00 is $4.90 \approx (1 - 0.1) \cdot 5 + (0.1) \cdot 4$.

In general, for each triplet of demand drop $\Delta log Q$, elasticity ϵ, and dollar threshold d, the bias is[24]:

$$\theta = d \left(1 - e^{\Delta log Q / \epsilon} \right) \approx -d \cdot \Delta log Q / \epsilon \tag{7}$$

The first estimation approach relies on this specification, where I estimate the demand drops for each product-dollar pair, and complement it with the elasticity from the IV specification in Equation 6. To get the

---

[22]The specification is $log Q_{ist} = \eta (1 - \theta) P_{ist} + \eta \theta \lfloor P_{ist} \rfloor + \alpha log P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu'_{i,on.sale(ist)} + \nu_{i,cents.digit(ist)} + e_{ist}$. Errors are clustered at the store level. Standard errors on θ are calculated using the delta method.

[23]The above single regression takes almost two weeks to run because of the about 470,000 fixed effects. For the NLLS method, it is unpractical for a numerical algorithm to be able to handle so many parameters.

[24]This is an approximation of $\Delta log Q = \epsilon (log ((1 - \theta) d + \theta (d - 1)) - log (d)) = \epsilon log \left(1 - \frac{\theta}{d} \right) \approx -\epsilon \frac{\theta}{d}$ for $\theta/d \ll 1$.

demand drops I estimate for each product

$$logQ_{ist} = \sum_{\{d=dollar-digit\}} (\epsilon_{id}logP_{idst} + \kappa_{id}) +$$

$$\alpha_i logP_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu_{i,on-sale(ist)} + \mu'_{i,spell.length(ist)} + \nu_{i,last-digit(ist)} + e_{ist}$$

clustering standard errors by store. Then I calculate drops for each product at each dollar digit x as

$$\Delta logQ_x = (\hat{e}_{x-1}log(x) + \hat{\kappa}_{x-1}) - (\hat{e}_x log(x) + \hat{\kappa}_x)$$

where I calculate standard errors using the delta method. As shown in Panel A of Table 2, I find an average drop of 9.0 percentage points (median of 8.0), where 86.4% of drops (keeping the most precisely estimated drop per product, in case there is more than one) are positive, and 54.5% of drops are significantly positive at the 5% level (compare with 2.8% significantly negative). The drops are a product-level econometric equivalent of the residuals figure, showing that this phenomenon is occurring product by product and not only at the aggregate.

Perhaps finding positive drops is an artifact of the data. For that reason I am running placebo tests where I am estimating the drop size in other cutoff points. For example, instead of assuming that people become fully aware of the exact price at round numbers (00-ending prices), I assume they do so at other points (e.g., at 50-ending prices, so they round to the nearest dollar). In that case, if taking 49 as the cutoff instead of 99, almost exactly half of drops are positive (48.9%), which we'd expect if this is happening at random. The estimation procedure still leads to 14.3% of drops being statistically significantly positive at the 5% level (and 19% significantly negative).

The drops, then, provide a reduced-form measure of the effect of a left-digit change. This is similar to Anderson and Simester (2003) or Macé (2012), who add a dummy for any 9-ending price. This exercise adds to that existing knowledge by doing so for many products over a large set of stores and chains. Further, combining with elasticity estimates gives a structurally meaningful measure – the left-digit bias is the ignored share of the last digits. This estimation approach, relying on Equation 7, has the advantage that the vertical drops are unbiased estimates even if there is price-ending sorting due to the bias, as long as there is no systematic sorting within product to price-endings by demand shocks. The elasticity estimates are unbiased to the extent that controls and instrumental variables solve endogeneity issues. Therefore, the left-digit bias estimates are arguably unbiased.

A histogram of estimates is shown in Figure 4. Table 2 shows that the mean parameter is 0.26, with interquartile range of 0.07 to 0.39. For the figure, I am highlighting in darker shade the estimates where the estimated drops is in the top three quartiles of precision. As can be seen in the figure, imprecise estimates generate most of the outliers. A CDF of all estimates is shown as the solid purple curve in Appendix Figure A.7.

4.2.2 Non-linear Least Squares

An alternative to the sufficient statistic approach is to estimate the model via non-linear least squares. The advantage of this approach is that it is a direct specification of the model. The disadvantage is that it relies on numerical optimization and is not run with instrumental variables. Namely, instead of regressing log-quantity on log-price, I regress log-quantity on log-*perceived*-price, and estimate the elasticity and θ simultaneously:

$$logQ_{ist} = \epsilon_i log\left(\left(1-\theta\right)P_{ist} + \theta\lfloor P_{ist}\rfloor\right) + \tag{8}$$
$$\alpha_i logP_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu_{i,on.sale(ist)} + \mu'_{i,spell(ist)} + \nu_{i,cents.digit(ist)} + e_{ist}$$

To estimate it I code the analytical gradient and hessian, and minimize the sum of squared errors with a numerical optimizer. I get the initial guess from running an OLS without the bias, and guessing a θ of 0.2 (I get the same results for different initial values of θ). The identifying assumption being that variation in prices and in price-endings is not driven by changes in elasticity. I am relying on variation in prices within and between dollar digits, absorbing the variation across years, month of year, stores, sale effects, and last digits with fixed effects. In the NLLS I do not instrument with other prices, and to get standard errors I use cluster bootstrap, sampling different stores with replacement. The CDF of product level estimates is the dashed blue curve in Appendix Figure A.7, which is very similar to the estimates coming from the drops. As shown in Table 2, the mean bias is 0.24 and the median is 0.18.

4.2.3 Semi-elasticity

A third method is to assume a log-linear structure instead of a log-log estimation. This is a further test of the way left-digit is modeled, since we should get the same estimates under a different demand structure. It has the key advantage that the bias can be estimated linearly in a simple OLS. That is, instead of the standard log-linear structure of $logQ \sim P$, I regress $logQ$ on $P + \lfloor P\rfloor$. That is, I run at the product level

$$logQ_{ist} = \eta_i\left(1-\theta_i\right)P_{ist} + \eta_i\theta_i\lfloor P_{ist}\rfloor + \tag{9}$$
$$\alpha_i P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu_{i,on.sale(ist)} + \mu'_{i,spell-length(ist)} + \nu_{i,cents.digit(ist)} + e_{ist}$$

where η is the semi-elasticity (the change in log quantity due to a unit change in prices), and the ratio of estimated coefficients on the price and the floor of the price gives the ratio of $\frac{1-\theta}{\theta}$. Reassuringly, the estimates from that approach are similar to the estimates from the NLLS approach, as can be seen in Figure A.7. The estimates have a correlation of 0.85 and a rank correlation of 0.91. As shown in Table 2, the mean bias is 0.20 with a median of 0.16.

The near-identical estimates across estimation techniques and demand structures imply that the left-digit bias parameter is robust to the assumed demand curve shape, and supports its interpretation as the neglect of lower digits.

Using a similar specification, I estimate the model with homogeneous semi-elasticity and left-digit bias across products (but with product level fixed effects and cross elasticities)[25]. I estimate a left-digit bias for these 1710 products of 0.1635 (0.096) and semi-elasticity of -0.3697 (0.0044). Since the mean price is \$3.50 it implies an average elasticity of -1.29.

---

[25]Specifically, I estimate

$$logQ_{ist} = \eta\left(1-\theta\right)P_{ist} + \eta\theta\lfloor P_{ist}\rfloor + \alpha_i P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu'_{i,on.sale(ist)} + \nu_{i,cents.digit(ist)} + e_{ist}$$

4.3 Exploring Bias Heterogeneity

The demand-side model makes no assumption about where the bias is coming from and what mental process it might represent. One immediate question to ask is what is correlated with the bias (see also Macé (2012)). The results provide field evidence of relative thinking that co-lives with a form of deliberate inattention. Because this is not at the core of the paper, I defer the analysis and discussion to Appendix E. In a nutshell, utilizing the Nielsen Consumer Panel (HMS) to gain characteristics of each product and the people who consume it, I find that the bias is increasing with the price, decreasing with the number of units a household purchases in a shopping trip, and is not significantly correlated with the purchasing clientele's education or income.

5 Firm Pricing

The demand side estimation corroborates the modeling of consumer behavior. Taking that as given, I turn to investigate the firm response. I document that qualitatively, firms act as if they respond to left-digit bias according to the model. The motivation for studying 99-prices is stemming from their prevalence, but the point is that other predictions of the model are also corroborated.

Since the firm seems to respond to the bias according to the model, I conduct an "as if" exercise, estimating the perceived left-digit bias from the pricing behavior. I find large gaps between the demand-side estimated bias and the supply-side estimated perceived bias.

5.1 Qualitative Predictions

As analyzed in section 2, the model prescribes three key qualitative predictions for pricing patterns of firms facing a demand structure such as the one estimated in section 4. The model predicts for optimizing firms: (1) excess mass at 99, (2) which is coming from a region of missing prices with low price-endings, (3) and increasing with the dollar digit.

The data selection process, motivated by the need to have variation in prices for the demand side estimation is problematic. Basically, I am selecting on the products that exhibit limited 99-pricing, so I am effectively forcing a lower bound on the mass at 99. Note that since the left-digit bias is estimated on the same set of products, this exercise is internally consistent. But, to be conservative and alleviate any concerns, I am also doing robustness tests with the entire sample of popular products without exclusions based on price distribution.

Excess mass   Figure 5 shows the price-ending histograms for the regular prices in the retailer data and the long price spells in the Nielsen data (at least 6 weeks). The reason I restrict attention to regular, or likely regular, prices, is that as shown above promotions techniques actually might work to increase demand better with round prices (see also Anderson and Simester (2003)). This is possibly because such pricing startegites, such as quantity discounts ("2 for $6"), tap on different psychological forces such as price fluency and suggestive quantities. These forces are beyond the model. Indeed, the model seems to fit best for the regular prices as in Figure 3(b).

High shares of prices indeed end with 99, 27% in the retailer data, and 34% in the Nielsen data. Also, almost all prices end with 9 (85% and 86% percent)[26]. This latter finding also implies that there is further bias ignoring the last digit of the price, hence leading to so many 9s. This effect is probably true, but impossible to estimate, exactly because there is lack of variation in the last digits, and hence the drops in demand at the dime thresholds can not be identified. However, I do robustness tests with additional bias regarding the last digits (that is, assume that with some measure θ_2 consumers see the price as the dollar and dime digits but ignore the cents).

Missing prices   Figure 5 also shows that for both the retailer and the Nielsen data, there are few prices ending with anything lower than 19. This finding has not been described, predicted, or explored in prior papers. Meaning, this finding provides further support for the model, especially because it is an unanticipated pattern and prediction.

It also does not seem that retailers price at 99 out of "convenience" or as if they perceive there is an inherent higher demand at 99 versus any other price (e.g., a "blip" in demand as discussed above). Such behaviors are manifested in symmetric missing prices, which is not the case here. In a related paper, Dube et al. (2017) study wage offer distributions. If employees are left-digit biased and want higher wages, then wages should be set at round numbers. The authors find excess mass at 0-ending wages, but missing prices are symmetric. They use the symmetry to argue against left-digit bias in their setting.

Next-lowest price by first digit   Figure 5 masks the heterogeneity of how price endings differ across dollar digits. The model predicts that the next-lowest price should increase with the dollar digit, keeping the bias and elasticity constant. If, as we have seen above, the bias is increasing with the price, this prediction intensifies. The share of 99-ending prices is not clearly predicted, since it relies on the distribution of costs.

In order to examine how price-endings change across dollar digits, I display the left tail of a by-dollar CDF of price endings in Figure 6. Darker lines are higher dollar digits. Indeed, Figure 6 shows that the distribution of price-endings is shifting to the right, meaning that for higher dollar digits there are less low price endings.

To recap, firm behavior is qualitatively consistent with the model. Prices bunch at 99, there are missing prices with low price endings, and the mass of low-ending prices is shifting to the right as the dollar digit increase.

5.2   Quantitative Predictions

But does the firm behavior align with the predictions of the model given the estimated bias and elasticities from the demand side estimation in Section 4?

Recall from Proposition 2 that next-lowest prices are decreasing as demand is more elastic and as the bias is weaker. Therefore, I take the low-end estimated bias level, of 0.16, and the mean elasticity from the IV specification of -1.5[27]. For these levels, the predictions are that next-lowest prices are 99. In that case there is no ambiguity, *all* prices should end at 99 regardless of the cost distribution.

Keeping in mind that firms do respond to the bias, qualitatively, I take an "as if" approach to infer the perceived left-digit bias levels that rationalize the retailer pricing behavior. I assume that the retailer is

---

[26]In the Nielsen data, the next most popular price endings are 0 (4.2%), 8 (3.7%), and 5 (2.7%). For the retailer, these are 5 (5.5%), 4 (1.9%) and 0 (1.6%).

[27]Since I am comparing the *regular* prices distribution to the estimates using all prices, it is worth noting that if I estimate the bias parameters using observations of regular prices only (with NLLS), I get slightly higher estimates, with median bias of 0.23.

pricing as described by the model, according to what they *perceive* to be the price elasticity and left-digit bias. In other words, in Section 4 I estimated the elasticity and bias that rationalize demand behavior, and here I estimate the elasticity and bias that rationalize supply behavior.

The first simple alternative is that elasticities are estimated with bias, but in reality (known to the firms) demand is more elastic. However, elasticity alone does not help. Even for extremely elastic demand (e.g. -15, compare with the most extreme estimates from Hausman (1996) of -3.17 or Nevo (2001) of -4.25), the next-lowest prices and shares of 99 are much higher than observed. But fundamentally, to the extent that the model is reasonable and demand estimation is proper, we should doubt extremely different perceived elasticity argument since the price elasticity is one of the key parameters a pricing firm cares about, so large discrepancies between the estimated and perceived numbers are unlikely. This might not be the case for the bias parameter since I am not aware of a portable way to estimate it and use it for pricing (which is what I try to do here). Indeed, if we assume that the estimated elasticity is correct but the bias is perceived to be much lower, then the pricing patterns can be rationalized quite well.

Next, instead of calibration I conduct proper estimation. The comparative statics effects of the left-digit bias and price elasticity allow for identification of the price elasticity and bias solely from the price distribution under the assumption of constant price elasticity and some parametric cost distribution. This is of methodological interest, since it allows for *price elasticity identification just from the price distribution even without knowledge of costs or quantities*. Absent the bias, this is impossible to do. Price data are becoming more and more available (e.g, Cavallo and Rigobon (2016); Cavallo (2016)), and this is a rare case in which a behavioral bias is allowing for estimation rather than adding a degree of freedom.

The idea behind the identification is that left-digit bias creates discontinuities, thus making frictions in pricing that can be exploited for identification. In a sense, this is the dual of the notch estimation in the taxation literature (see Saez (2010); Kleven and Waseem (2013))[28]. However, this estimation approach relies on measuring with precision an ill-behaved statistic which is the next-lowest price. An extremum moment is highly sensitive to measurement error, and even small heterogeneity can lead to disproportional large effects on the estimands[29]. In addition, the price distribution is more discrete than just masses at 99, as almost all prices end with 9. If this is driven by added left-digit bias to the cent digit, different parameters would translate to the same 9-ending next-lowest price, tampering with the variation needed for identification.

From a practical standpoint, this is also a data-intensive method. It hinges on the assumption that the cost distribution is smooth for regular prices. While this is probably true aggregating over many products, it is less likely to hold for a single item. Further, if a product is priced identically across stores (as is true in the data, especially for regular prices), increasing the number of stores does not affect precision.

To satisfy the need for a smooth cost distribution, I am aggregating over products, assuming a constant left-digit bias and elasticity. Which leads me to the next caveat - the assumption of homogeneous parameters. Heterogeneity of the bias parameter can of course lead to low-ending prices, and is definitely plausible, as shown in Section 4. However, identification with heterogeneous bias relies far more heavily on the shape of the cost distribution, which is an arduous requirement to infer from the data. I argue that since there is such a stark difference in predictions on aggregate and product-by-product (there are only 11 out of 3859

---

[28]In Public Economics the idea is that non-linear tax schedule creates incentives for tax payers to change their reported income (or labor supply) and bunch at kinks or notches of the tax schedule. The duality can be seen as follows: budget set of taxes is like the demand (elasticity); the notch given by the tax schedule is like the bias parameter; the utility function is like a profit function; and the latent ability distribution is like the cost distribution - estimated from the observed income distribution like the observed prices outside of the bunching and missing prices.

[29]Kleven and Waseem (2013) face a similar issue where no one should have income in dominate regions. They interpret it as frictions, and implicitly assume that frictions are random. The approach I follow is implicitly taking a similar stance, since I put equal weight on many price moments.

popular products whose prices end solely by 99), assuming a single left-digit bias parameter is a reasonable first step. In addition, in the next section I explore the implications of heterogeneous parameters on pricing and profits, and find them to be small.

Given all the caveats, this is an exercise worth doing even if only for getting rough estimates. To be less sensitive to the exact measure of missing prices, I will use the entire price distribution (binned at 9-ending prices) and a method-of-moments approach[30].

Let S_p be the share of prices set at a price p. I use minimum-distance estimation to match the empirical price densities, $\{S_p\}$, to the predicted price densities, $\{\hat{S}_p\}$ (where p belongs to a grand price vector $\mathcal{P}$). For example, take $\mathcal{P} = \{2.29, 2.39, ..., 3.69\}$. The algorithm is as follows: (1) Fit a polynomial to the empirical price distribution using prices that are not bunching at 99 or missing due to the bias, call that price distribution $\hat{F}_p$. Then, for each pair of elasticity ϵ and left-digit bias θ: (2) take all relevant 99-ending prices, $\{q_i\}$ and calculate the matching next-lowest prices $\{P_{q_i}\}$ using Equation 5 (for example, say $\{q_i\} = \{1.99, 2.99\}$ and that $\{P_{q_i}\} = \{2.51, 3.53\}$). (3) expand the vector of prices to include $\{q_i\}$, $\{P_{q_i}\}$, and all 9-ending prices between each pair of P_{q_i} and q_{i+1}, giving a vector of prices $\mathcal{P}'$ (e.g. $\mathcal{P}' = \{1.99, 2.51, 2.59, 2.69, ..., 2.99, 3.53, 3.59, 3.69\}$). (4) Given θ and ϵ calculate for each $p_i \in \mathcal{P}'$ the cost $\underline{c}_{p_i}$, such that p_i is the profit-maximizing price. By construction this is given by the first order condition as $\underline{c}_{p_i} = p_i \frac{1+\epsilon}{\epsilon} + \frac{\theta}{1-\theta} \frac{\lfloor p_i \rfloor}{\epsilon}$. (5) Then for each price, the share of observations at that price are $\hat{S}_{p_i} = \hat{F}_c\left(\underline{c}_{p_{i+1}}\right) - \hat{F}_c\left(\underline{c}_{p_i}\right)$. And we get $\hat{F}_c$ using the following model-induced identity:

$$
\hat{F}_c(\underline{c}_p) = Pr(c \leq \underline{c}_p) = Pr\left(p\frac{1+\epsilon}{\epsilon} + \frac{\theta}{1-\theta} \frac{\lfloor p \rfloor}{\epsilon} \leq \underline{c}_p \right)
$$
$$
= Pr\left(p \leq \underline{c}_p \frac{\epsilon}{1+\epsilon} - \frac{\theta}{1-\theta} \frac{\lfloor p \rfloor}{1+\epsilon} \right)
$$
$$
= \hat{F}_p\left(\underline{c}_p \frac{\epsilon}{1+\epsilon} - \frac{\theta}{1-\theta} \frac{\lfloor p \rfloor}{1+\epsilon} \right)
$$

(6) expand $\mathcal{P}'$ to $\mathcal{P}$ by adding the missing prices, between q_i and P_{q_i}, and assign them predicted zero shares[31].

The above procedure generates the predicted price moments $\hat{S}_{p_i}$ as a function of elasticity ϵ and left-digit bias θ. I minimize the distance between predicted moments and actual moments. Since I am also estimating the densities for prices outside of the missing prices and 99s, I am excluding these moments from the minimization problem.

This procedure might be somewhat sensitive to which moments are chosen, what moments are excluded from the fit estimation, and the degree of the polynomial fit. I therefore take 32 combinations of estimation settings[32]. For example, in one of these settings moments are between \$1.29 and \$5.89, using all prices with price endings between 29 and 89 for the $\hat{F}_p$ fit, with a polynomial of degree 3. I minimzing the distance of actual moments from predicted with price endings strictly higher than 89 and lower than 29 (that is 99, 09, and 19), since these are the prices that were not used to predict the price distribution[33]. Then, for each

---

[30]Maximum likelihood will give 0 bias because any low price-ending is a zero-probability event.

[31]relabel non-9-ending next lowest prices as the lower 9-ending price, without changing the share. For example, 2.51 will be relabeled as 2.49 with a share $\hat{S}_{2.49} = \hat{F}_c\left(\underline{c}_{2.59}\right) - \hat{F}_c\left(\underline{c}_{2.51}\right)$

[32]I take all combinations of

$\{bottom - cents\} \times \{top - cents\} \times \{lowest - dollar\} \times \{highest - dollar\} \times \{polynomial - degree\} = \{0.29, 0.59\} \times \{0.79, 0.89\} \times \{\$1, \$2\} \times \{\$5, \$6\} \times \{3,$

[33]Appendix Figure A.11 shows an example of the actual moments, highlighting which were excluded from the fit to create the counterfactual distribution in light yellow. The purple bars are the shares that are being matched by the predicted moments. The

setting I run 300 cluster-bootstraps at the product level.

Table 3 presents the estimation results showing the estimated perceived left-digit bias and price elasticity for the main sample, using regular prices of the 1710 products. I am guessing 50 initial values of bias and elasticity and choose the estimation with the best fit. To get confidence intervals I run additional thousand cluster-bootstraps at the product level (each with 50 initial guesses). I am estimating homogeneous parameters and find a mean perceived bias of $\hat{\theta} = 0.023$ and elasticity of $\hat{\epsilon} = -3.35$. I can reject a perceived left-digit bias larger than 0.043, but can not reject that the price elasticity is -1.5.

If I restrict the elasticity to be -1.5 and only estimate the bias I get tighter estimates (with similar goodness-of-fit) for the same bootstrapped sample of $\hat{\theta} = 0.0047$. The similar goodness-of-fit between the unrestricted and restricted versions means, as predicted by the comparative statics results, that there is weak identification separating between the elasticity and left-digit bias parameters. However, the firm is underestimating the bias at least by a factor of 3, and more likely by an order of magnitude. As a robustness, a break-down of the estimates by specific setting with 95% confidence intervals per each setting is plotted in Figure 7. The figure shows that while point estimates are sensitive to the moments choice and the assumed shape of the cost distribution (via the polynomial fit), confidence intervals are stable.

The two rightmost columns of Table 3 are showing the results using the alternative sample of popular products, namely those with less variation in prices (out of the 2149 products, 31 have shares of 99-ending prices above 75%, and the rest have most of their prices concentrated within a single dollar digit). The same pattern arises, with very similar estimates for both samples, and similar patterns of a flat goodness-of-fit between the restricted and unrestricted specifications. This is useful since there might be a worry that by choosing the 1710 products that exhibit less variation in prices leads to a mechanical finding of lower perceived bias (even though the demand side estimates should be consistent with their pricing if the bias is perceived correctly). Finding similarly low perceived bias for the products with more 99s and less between dollars variation adds to the robustness of the results.

This method can be used to estimate the price elasticity from a price distribution in other settings, especially if one is willing to take a stance regarding the left-digit bias. Meaning, if the researcher does not have access to quantities sold nor costs but observe multiple prices for similar products with fluctuating costs, they can estimate what the firm perceives to be the price elasticity. Appendix Figure shows the estimated elasticities for the various specifications.

6   Counterfactuals

The firm perceives the bias as smaller than it is. But perhaps this is not a costly mistake. Using the estimates from the demand side as "the truth", I compare profits under different pricing strategies: if the firm is ignoring the bias, fully optimizing, or pricing by the perceived bias as estimated in Section 5.

To be precise, I do the following: I am simulating a cost distribution by taking the actual price distribution scaled down by a constant markup. I calculate the price for each triplet of cost, the firm perceived price elasticity $\hat{\epsilon}$, and bias $\hat{\theta}$ (assuming elasticity is perceived correctly $\hat{\epsilon} = \epsilon$). Given the true parameters of the price elasticity ϵ, and θ, I calculate the profit from each tuple and then integrate. That is, I calculate

$$\Pi = \int D\left(p\left(c;\epsilon,\hat{\theta}\right);\epsilon,\theta\right) \cdot \left(p\left(c;\epsilon,\hat{\theta}\right) - c\right) dF\left(c,\epsilon,\theta,\hat{\theta}\right)$$

predicted moments are shown in semi-transparent bordered bars.

25

The formula reflects that pricing is governed by the perceived bias, while demand is governed by actual left-digit bias.

The results are shown in Table 4, where I present the gross profits lost by pricing according to $\hat{\theta}$ versus the profits if the firm were to price according to θ. I assume different distributions of the elasticity, left-digit bias, and perceived left-digit bias parameters. Panel A shows that if the firm were to ignore the bias altogether, that is price as if $\hat{\theta} = 0$. Panel B shows the loss if the perceived bias is as estimated, lower than the actual but not ignored.

I take the point estimates from the demand and supply side as benchmark, and estimate that the retailer is losing 2.7% of profits versus the optimum, out of a possible loss of 4.3% if were to ignore the bias. If I assume that the estimation is flawed, and the bias is actually lower, at 0.1, the elasticity is on the high end from the perceived estimation, at -4, and the perceived bias is also higher at 0.03, I still find losses of 1.1% of gross profits. I then explore the effect of heterogeneity of each of the parameters, and it does not carry a large impact on the results. Finally, I also estimate a case of a 2-product monopolist with an elasticity of substitution of 0.1 between them. For the particular distribution the potential loss is 3%, and the perceived bias leads to a loss of 1.7%.

In Appendix Table A.3 I add two robustness tests for each of the scenarios. In the first I assume that consumers also ignore the last digit with $\theta_2 = 0.25$. It bears very small effect on the results. The second robustness is more fundamental. I assume that the bias is such that when consumers ignore the lower digits they replace them with .99 instead of .00. It decreases the losses to be 1%-2% across scenarios.

Across scenarios and assumptions I estimate that the retailer may lose 1.1%-3.2% of gross profits. If, as in Montgomery (1997), operating margins are about 12%, it means that the firm is losing 9%-27% of operating profits by underestimating the bias.

The losses are coming from pricing at dominated regions. That is, the firm is pricing below the next-lowest prices and not at the lower 99-ending price. Pricing in a lower 99-ending price is beneficial exactly because compared to pricing right above it, consumers purchase higher quantities by more than the pure price elasticity effects. Note, also, that as Appendix Figure A.9 shows, the overall price level actually does not change much, and if anything is somewhat lower. This result is driven by 99-ending prices being mostly drawn from higher prices. Whether consumers replace the price with the floor, ceiling, or any other arbitrary component, does not affect this result.

7 Conclusions

In this paper I show that a model of left-digit biased consumers and optimizing firms is a good description of retail data. Consumers exhibit substantial left-digit bias in everyday choices while shopping at supermarkets. The supermarkets, in response, are using 99-ending prices and avoid some low-ending prices to increase demand. However, retailers also seem to underestimate the extent of left-digit bias. I find that firms do better than pricing as if there is no bias at all, but not half as good as possible. I estimate that they lose 1%-3% of gross profits, implying losing a significant share of operating profits. How can it be, that firms both respond to the bias, but stop short of fully optimizing?

Explanations can be divided to two categories, where either the model is missing something fundamental, or the firms are making a consistent mistake.

First and foremost, the pricing model is very simple. It simplifies competitive forces to enter through the price elasticity. Pure price competition with left-digit biased consumers also leads to missing prices as shown in Appendix B, but I did not solve a full model that can be taken to the data. Insofar as elasticity is a good approximation for the strength of competition it does reduce the ranges of missing prices. But, as I argued above, high elasticity by itself is not a plausible resolution to the discrepancy. Next, consider the case of a multi-product monopolist, as retailers are. Keeping same-price elasticity and left-digit bias fixed, if differentiated products are substitutes, it leads to *higher* next-lowest prices. This result holds whether the other product is more or less elastic, and whether its cost is higher or lower (see Appendix C). The conclusion, in any case, is that adding cross elasticities to the counterfactual analysis is likely to make the discrepancy stronger.

Another explanation is that the bias level might be endogenous to the price distribution. Meaning, if all prices were to end at 99 as the estimation suggests, then the bias would have become much lower. This is a sensible argument, but the needed effects on left-digit bias seem unlikely strong. For example, assume that left-digit bias is decreasing in the next-lowest price. Say that it is 0.15 if the next-lowest price is $2.19. In order for $2.19 to be an optimal next-lowest price, left-digit bias should decrease to the degree that is eliminated completely if the lowest price used is instead $2.21[34]. This seems like an unlikely strong effect.

Another avenue is that the model is a reasonable approximation, but firms are wrong. One possibility is that firms are following a heuristic ("use 99-prices") and so when the "absent-bias" price they come up with is close to 99 they round to 99[35]. The appeal of this story is that it makes intuitive sense that an absent-bias price of $2.00 should be adjusted to $1.99, but that it is less clear what a $2.50 should change to. But there is more to it, since not only is it more intuitive to make small changes in prices, but also the impact on profits is larger for the small changes. For example, consider two products with unit costs of $1 and $1.25, price elasticity of -2, and bias of 0.15. Absent the bias the product should be priced at $2 and $2.50, but with the bias both should be priced at $1.99. The gains from changing from a naive $2 to an optimal $1.99 are 15.5 percentage points in profits, while the gains from changing a naive $2.50 to an optimal $1.99 are a much lower, at 2.6 percentage points. So, given the counter-intuitive nature of big adjustments in price, and considering the relatively lower gains of big adjustments (masked in fluctuations in demand in real data), it seems reasonable that firms would stop shy of full adjustment. Making the full adjustment requires either "brave" and large scale experimentation with prices, or a rigorous analysis with a quantitative model.

A companion paper, Shlain (2018), analyzes a policy reform in Israel where supermarkets had to stop the common practice of pricing at 99-ending prices, and were only allowed to price "to-the-dime" (i.e, in 0-ending prices). Before the reform about 40%-50% of prices ended with 99, suggesting substantial levels of perceived and actual bias. Using the same model as in this paper, I show that for any reasonable level of left-digit bias, firms should have never round to a 00-ending price. However, all chains did so, and priced about 20% of products at 00-ending products after the reform. Within 6-12 months, they stopped. This finding is inline with the heuristic explanation, showing that firms respond to the bias in a limited "trial and error" sense, but without understanding the underlying model.

---

[34]To see that, note that if 2.19 is indeed the next-lowest price, it should satisfy optimality such that $\frac{\partial \Pi}{\partial p} = 0$, where I allow for θ to be a function of P. So, it is easy to see that the slope of bias to next-lowest price in that case is $\frac{\partial \theta}{\partial P} = \left(\epsilon \left(1 - \theta \right) + \frac{1}{P - P_1} \right) P - \theta$, which equals 8.6 for elasticity of -1.5, bias of 0.15 and P of 2.19.

[35]also "corroborated" by three anecdotal conversations I held with two mom-and-pop shop owners and a CEO of a large chain of convenience stores.

Figure 1: Illustration of the optimal price schedule
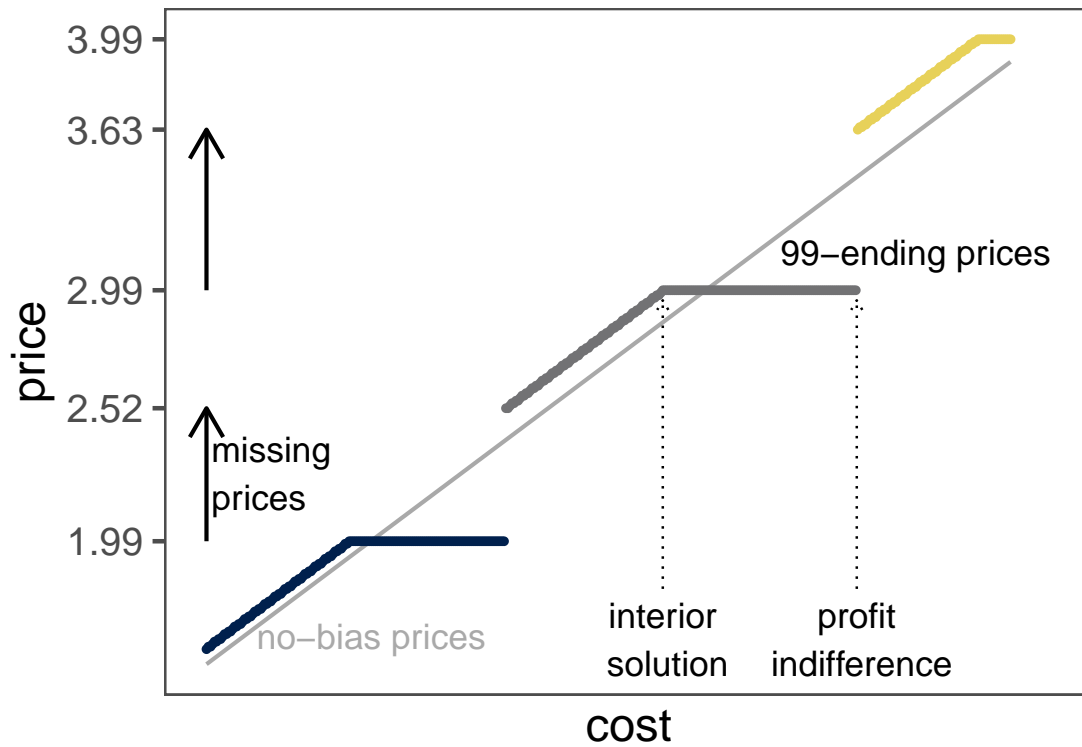
The thick lines represent optimal prices as function of cost, of a monopolist facing constant elasticity demand with left-digit biased consumers. The gray thin line represents the no-bias counterfactual where $p = c\frac{\epsilon}{\epsilon+1}$. The horizontal sections of the thick lines are the 99-ending prices, emerging for regions of costs. The lowest cost that translates to a 99-ending price is when this is the interior solution, and the highest cost is when there is profit indifference between 99 and the next-lowest-price. The resulting ranges of missing prices are highlighted by vertical arrows.

Figure 2: Illustration: Demand and residuals with and without left-digit bias

(a) Illustration: Demand with left-digit bias
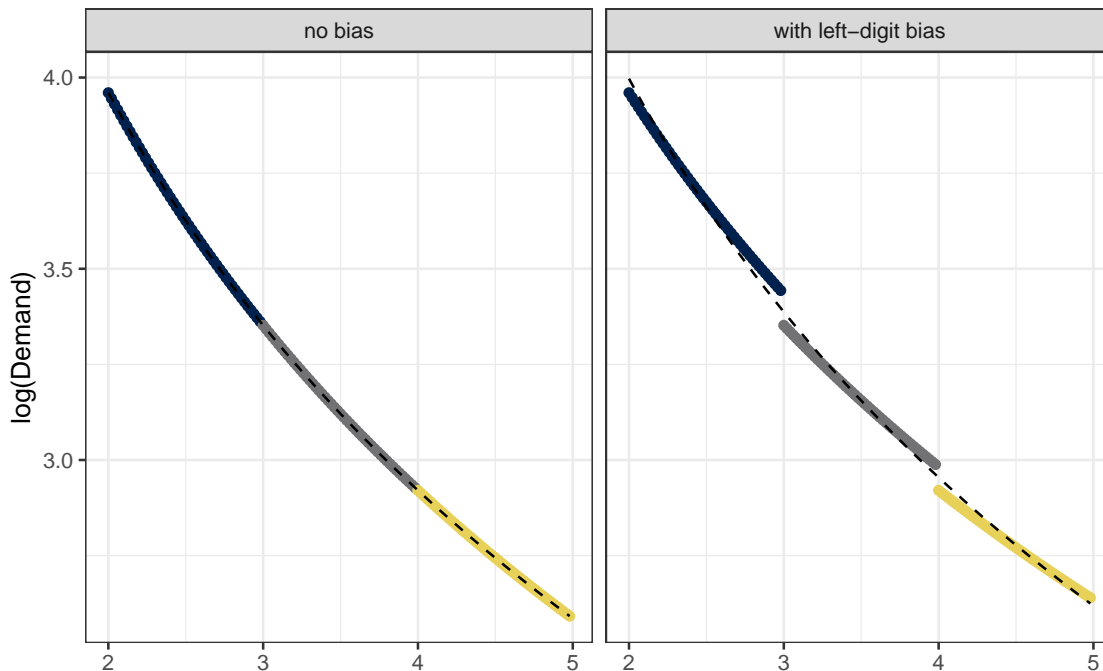
(b) Illustration: Residualized demand from log-log regression with left-digit-biased demand
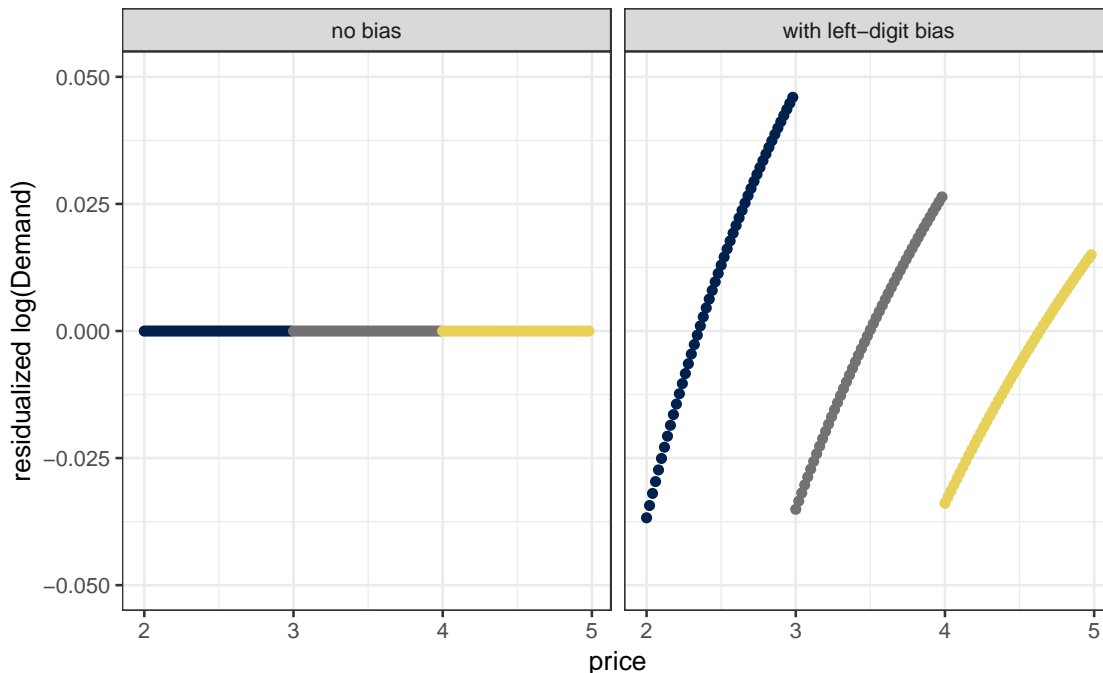
The figures illustrate simulated demand curves under a model of left-digit bias. The top figures show simulated log(quantity)-price demand curves where $logQ = 5 - 1.5log\left((1-\theta)\,p + \theta\lfloor p\rfloor\right)$. Different colors represent different dollar digits. The dashed gray line is the curve of best fit from a logQ~logP regression. The bottom figures show the residualized demand (actual minus predicted) by price. The right figures are for a case left-digit bias of $\theta = 0.16$, and the left figures are for the standard case of no bias, $\theta = 0$.

Figure 3: Sawtooth patterns of residualized demand by price

(a) Main sample, IV estimation

(b) Restricted sample of regular prices, OLS estimation
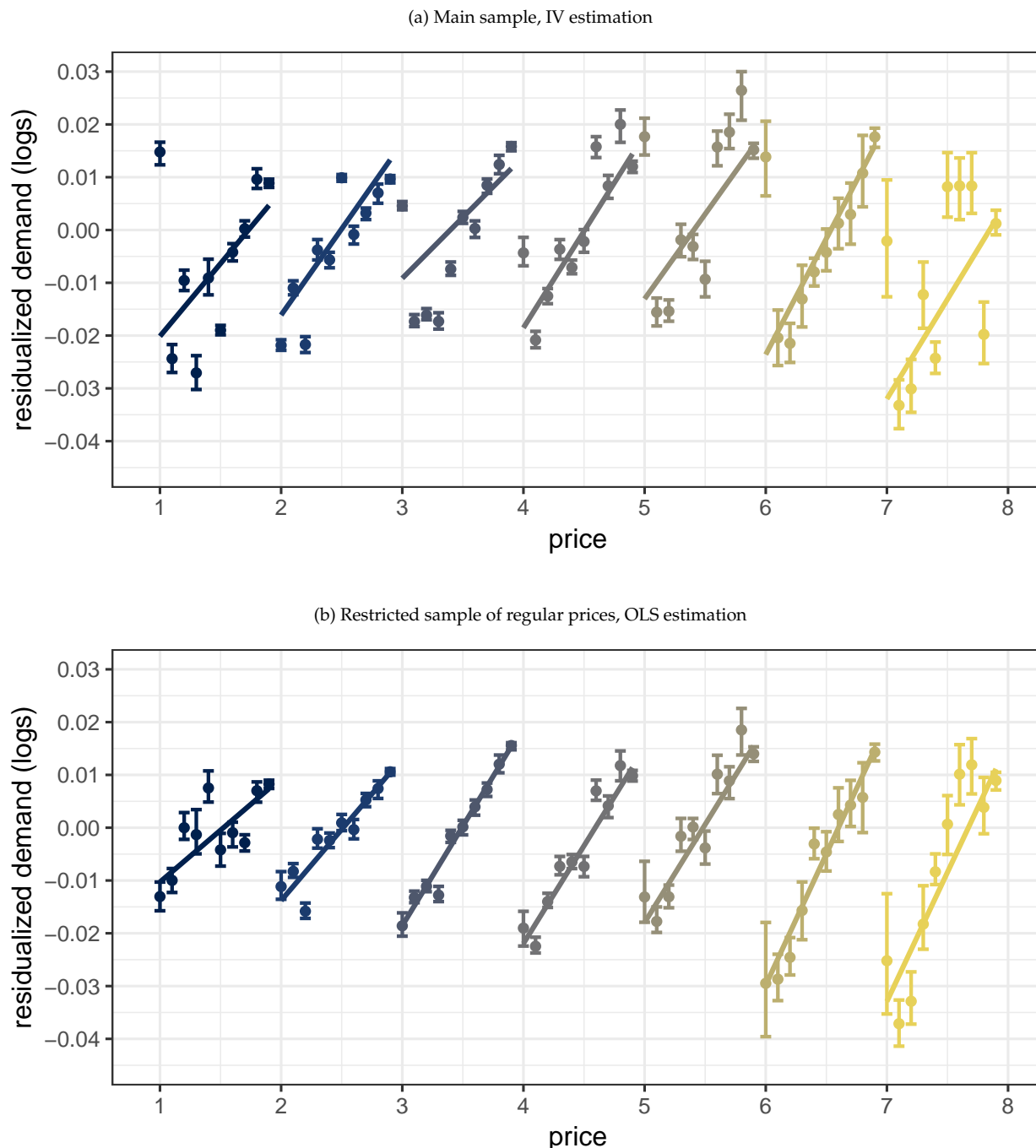
Net log-demand by price, residualizing product-level price elasticity, seasonality, cross elasticities, store fixed effects, and promotion effects. Regressions are conducted at the product level (1710 products, 248 stores, 177 weeks). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Plotted fixed effects estimated at 10-cent bins (e.g. $3.90-$3.99), and 95% confidence intervals are the result of 300 cluster-bootstraps by store.

The top panel is using data on all to-the-cent prices from a national retailer data (Column (3) in Table 1), and the regression is run with instrumental variables, as specified in equation 6 in Section 4.1. The bottom panel presents the results of an OLS constructed on regular prices only (column (4) of Table 1). Data is truncated for prices between $1 and $7.99 for clarity. For further details see Section 4.1.

Figure 4: Histogram of product-level left-digit bias estimates from drops in demand at the dollar-digit
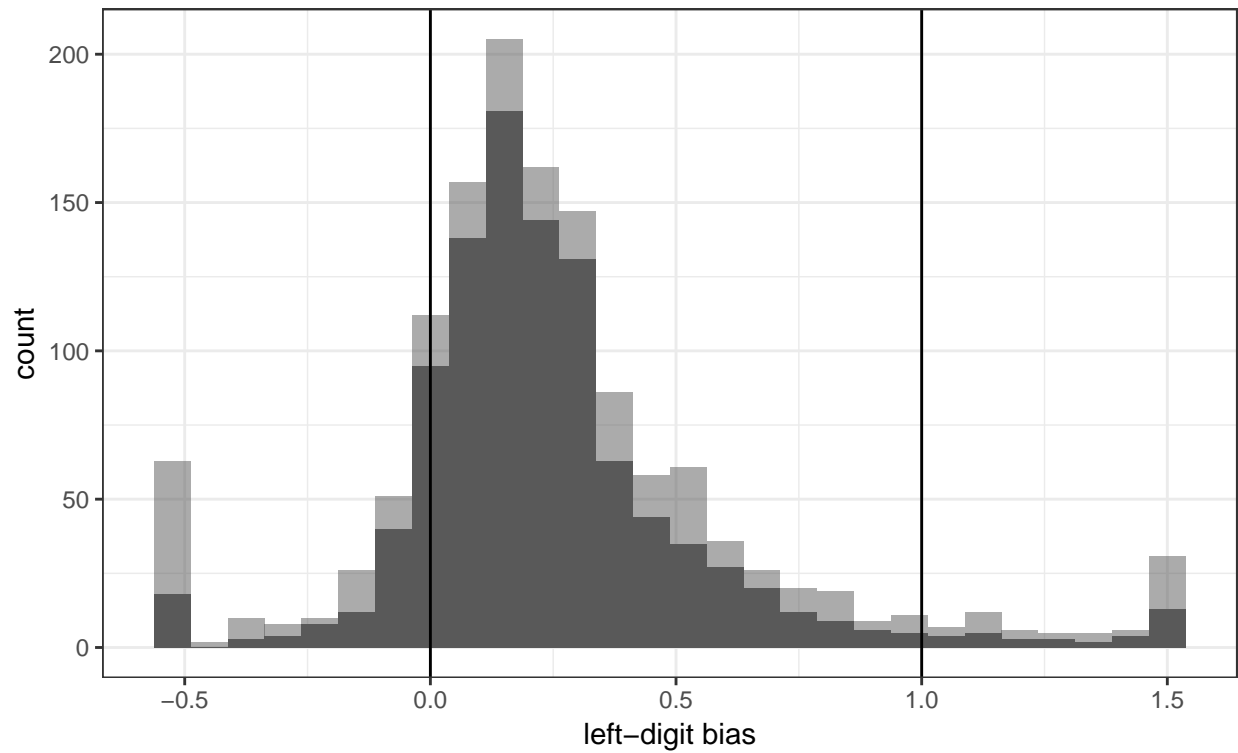
The figure shows the histogram of product-level estimates of left-digit bias, calculated from Equation 7, using estimated drops in demand and price elasticity estimated with instrumental variables. Highlighted by darker color are the 75% most precise estimates. Most impercise estimates are also outliers and vice versa.
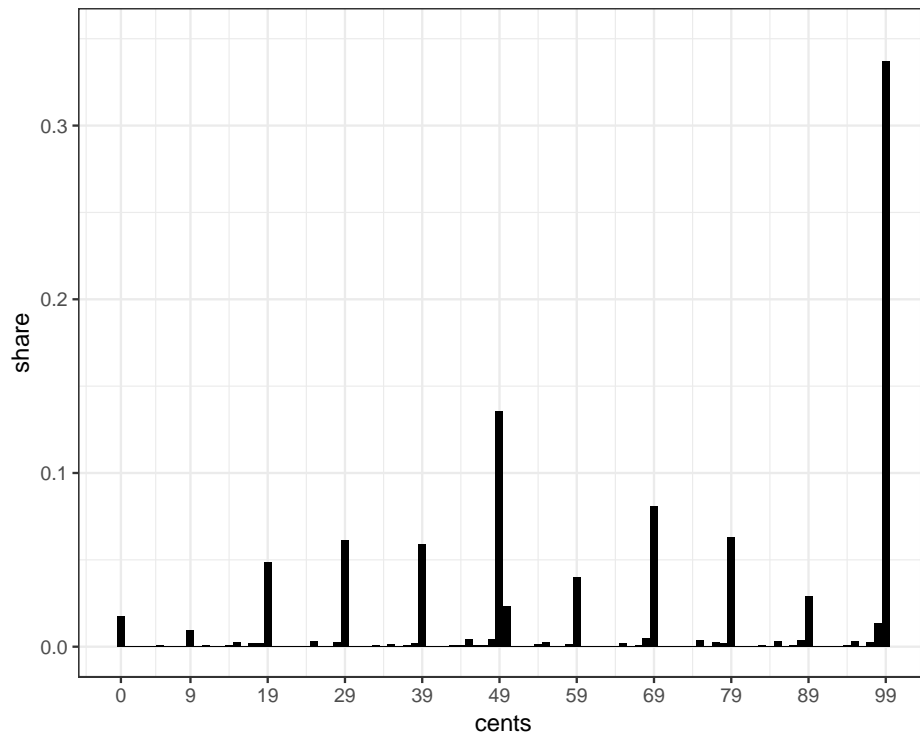
Figure 5: Pricing response: Price-endings distributions

(a) Retailer Data
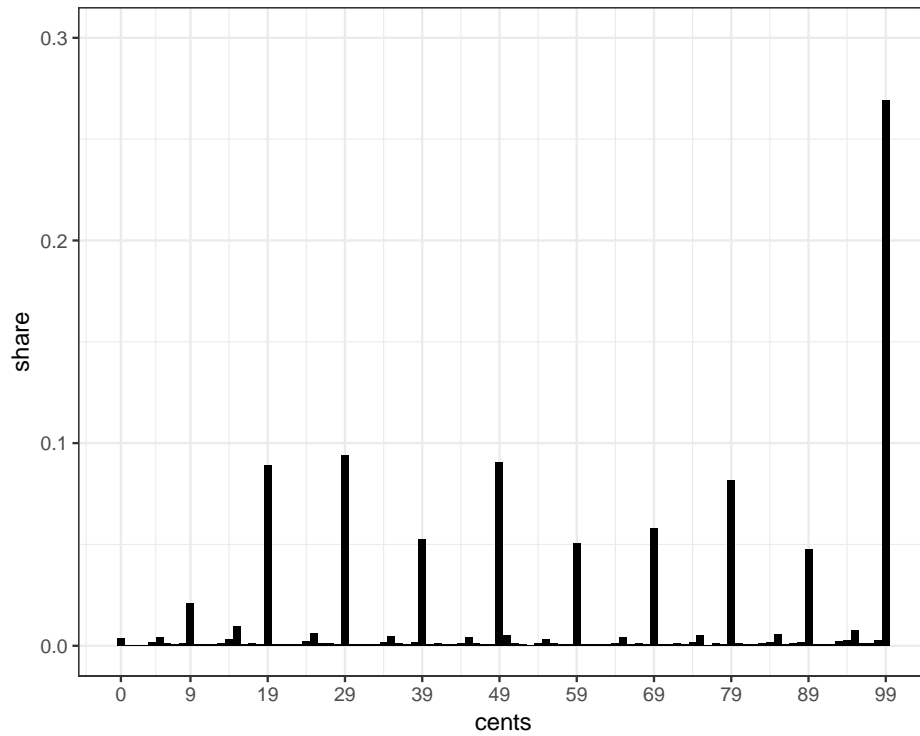
Source: 1710 products, regular prices, retailer.

(b) Nielsen Data

Source: 12 products, spells of at least 6 weeks, Nielsen.

The figures show the price-endings histograms of all regular prices in the retailer data (36M observations, top panel), and all prices that belong to spells of at least 6 weeks in the Nielsen data across more than 60 retailers (11M observations, bottom panel). Main patterns are excess mass at 9- and 99-ending prices, and missing prices at low price endings.

Figure 6: Price endings: left-tail of CDF by dollar digit in retailer data

(a) Left tail of CDF for regular prices of main sample

(b) Left tail of CDF for regular prices of full sample
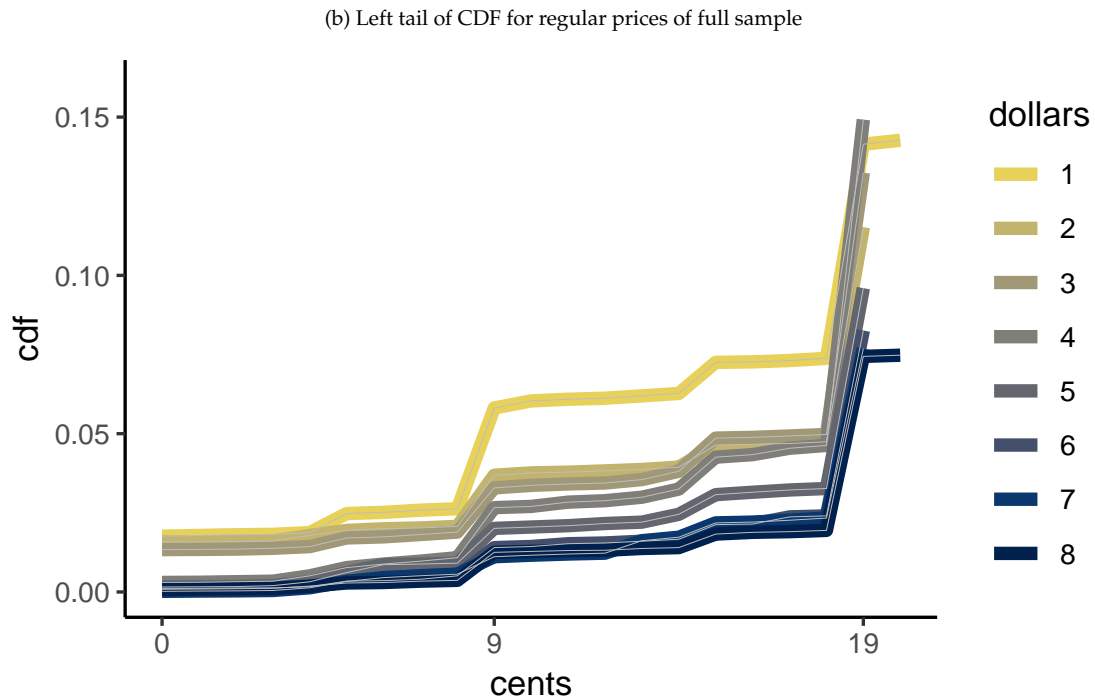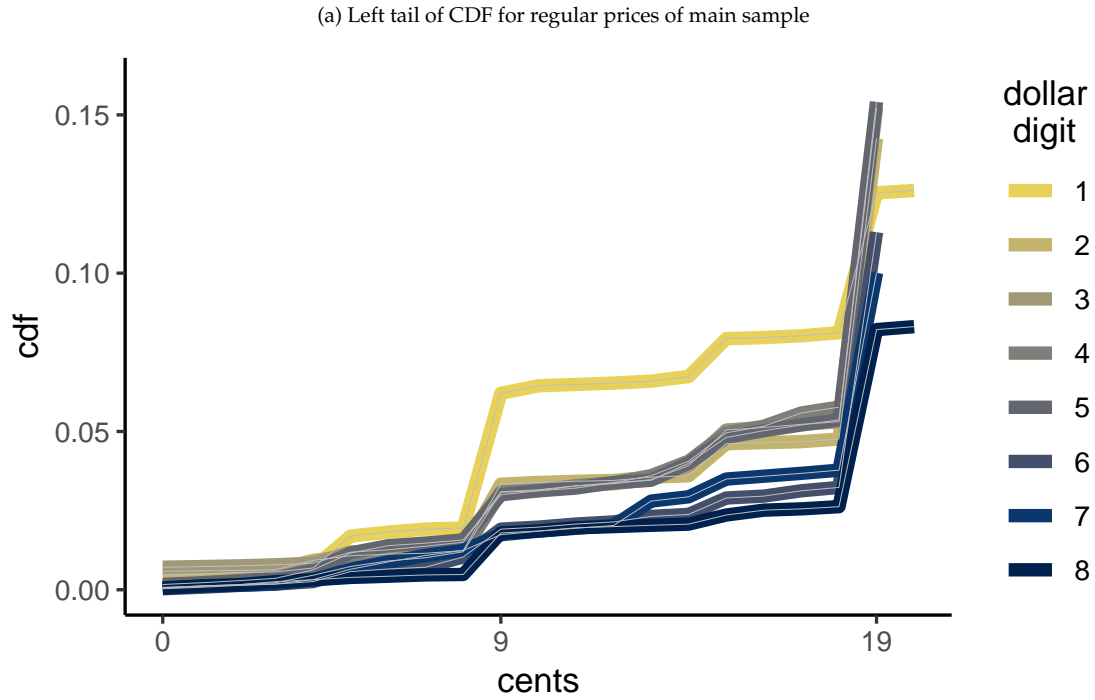
The figures show the left-tails of regular prices price-endings empirical CDFs, by dollar digit. The top panel is the main sample, while the bottom uses all regular prices in the data. Line color represents the first digit of the price, where $1+ is the lightest and $8+ is the darkest. The figures show that as prices increase, the firm is using less of lower price endings, qualitatively following the predictions from Corollary 1.

Figure 7: Estimated firm-perceived left-digit bias
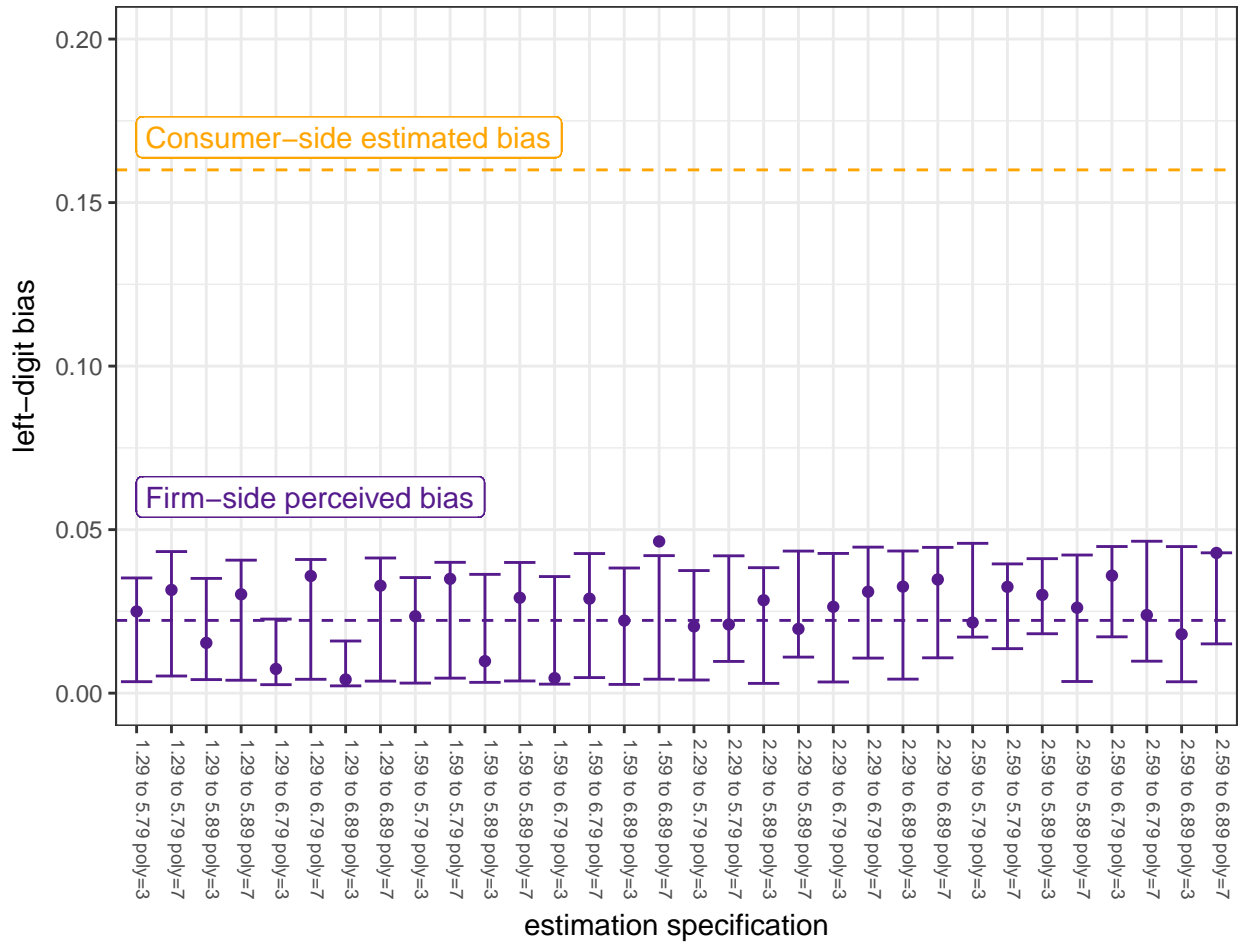
The figure shows the estimated perceived left-digit bias inferred from the price distribution. On the horizontal axis are different assumptions about the estimation procedure. Namely, the range of prices used for estimation, the excluded prices from the counterfactual price distribution fit, and the shape of that distribution via the polynomial degree. Each point is the estimated perceived left-digit bias for each specification, and the confidence intervals are the 2.5% and 97.5% estimates from 300 cluster-bootstraps of the moments at the product level (that is why some point estimates are outside of the confidence interval).

For all specifications, I can reject that the perceived bias is 0 or larger than 0.05. While point estimates are sensitive to the exact specification, bootstrapping the sample provide more similar confidence intervals.

Table 1: Summary Statistics and Data Selection

	Retailer Data				
	Full	1710 UPCs	Main sample	Regular Prices	Wholesale (costs)
	(1)	(2)	(3)	(4)	(5)
Panel A: Data Description					
Observations (Millions)	375.0	51.0	45.9	36.2	375.0
Products	74224	1710	1710	1710	74224
Stores	250	248	248	248	250
Weeks	177	177	177	177	177
Chains	1	1	1	1	1
Annual Revenue ($Millions)	2060	396	304	227	-
Annual Units Sold (Millions)	883	136	99	70	-
First Date	Jan-04	Jan-04	Jan-04	Jan-04	Jan-04
Last Date	May-07	May-07	May-07	May-07	May-07
Panel B: Pricing Descriptives					
Mean Price	3.10	3.42	3.50	3.66	2.14
Price SD	1.97	1.63	1.64	1.67	1.59
Share on-sale	0.36	0.29	0.22	0.00	
Share 99-ending	0.20	0.22	0.24	0.27	0.01
Share 9-ending	0.61	0.67	0.74	0.85	0.11
Share 9-ending (non-sale)	0.82	0.85	0.86	0.85	0.12
Share 0-ending	0.19	0.13	0.13	0.02	0.12
Share 0-ending (non-sale)	0.05	0.02	0.02	0.02	0.12

The table presents summary statistics for the retailer data. Column (1) is the full data set, after exclusion of by-weight products. Column (2) shows the sample of 1710 products chosen for analysis. Column (3) is the main sample used in the paper, excluding non-real prices (i.e., keeping only prices that are to-the-cent). Column (4) provides further information on the instances in which the price charged is the regular price and not a discounted one. Column (5) provides some descriptive statistics on the price endings of the wholesale prices the retailer pays for the same products as in column (1).

Table 2: Estimates of product-level drops and left-digit bias parameter

	Mean estimate	Number of estimates	Percentile 25th	median	75th
Panel A: Drops in demand (percentage points)					
(1) All drops	8.89	2518	1.79	8.22	15.5
			(4.74)	(8.33)	(3.99)
(2) Most precise drop per product	9.01	1351	3.26	8.04	13.8
			(3.22)	(3.51)	(6.18)
(3) Drops estimated from regular prices sample	8.71	1029	2.06	7.90	15.2
			(2.62)	(2.88)	(2.21)
Panel B: Left-digit Bias Estimates					
(4) Homogeneous, semi-elasticity model	0.164	1			
	(0.010)				
(5) Drops and IV elasticities	0.226	1351	0.065	0.202	0.388
(6) Non-linear least squares	0.239	1701	0.046	0.181	0.374
			(0.061)	(0.067)	(0.159)
(7) Semi-elasticity specification	0.200	1632	0.029	0.165	0.333
			(0.054)	(0.081)	(0.053)

Panel A shows the distribution of product level estimated drops-in-demand at dollar crossing, and panel B shows the distribution of left-digit bias estimates. In parentheses are standard errors clustered at the store level (asymptotic in rows 1,2,3,4, and 7; by cluster-bootsrap in row 6). The column titled number of estimates shows the number of estimates in each row. They differ from the number of products due to convergence problems, or lack of sufficient variation.

Panel A: Row (1) shows all estimated drops in demand, including multiple drops per products (for example if a product prices span $2.19-$4.79 then it will have a drop at $3.00 and at $4.00). Row (2) keeps the most precisely estimated drop per product. Row number (3) uses a restricted sample, where the drops are estimated from regular prices only.

Panel B: Row (4) provides the estimate from a homogeneous log-linear model, with a single parameter of left-digit bias and single semi-elasticity. Rows (5)-(7) show the product-by-product estimates, inferred by three different approaches: The sufficient statistic approach (Equation 7 in Section 4); estimated by non-linear least squares (Equation 8); and log-linear OLS (Equation 9), respectively. To calculate means I winsorize at $\theta = -0.5$ and 1.5. See Figure A.7 for the CDFs of the complete distributions.

Table 3: Estimated perceived parameters from firm pricing behavior

	Main Sample		Alternative Sample	
	Estimating Both	Fixing Elasticity	Estimating Both	Fixing Elasticity
Left-digit bias	0.023	0.0047	0.028	0.0041
	[0.004, 0.043]	[0.0040, 0.0069]	[0.006, 0.045]	[0.0024, 0.0056]
Elasticity	-3.35	-1.5	-4.15	-1.5
	[-1.40, -4.97]	[.]	[-1.91, -5.31]	[.]
Goodness of fit	0.00347	0.00355	0.00575	0.00580

The table shows the results of minimum distance estimation of homogeneous perceived left-digit bias and price elasticity using the regular prices distribution of 1710 products priced by the retailer (two left columns) and an alternative sample of 2149 popular products in the two rightmost columns. Confidence intervals in square brackets are the result of 300 cluster-bootstraps per 32 different specifications (total of 9632 runs per column) at the product level.

Table 4: Profits lost when the firm underestimates left-digit bias

	Assumed scenario			% profits lost pricing
	Left-digit bias θ	Price elasticity ϵ	Perceived bias $\hat{\theta}$	by perceived bias ($\hat{\theta}$ instead of θ)
Panel A: If the firm ignores the bias ($\hat{\theta} = 0$)				
Point estimates	0.16	-1.5	0	-4.3
Extreme estimates	0.10	-4.0	0	-4.4
Heterogeneous bias	$N(0.16, 0.05)$	-1.5	0	-4.9
Heterogeneous perceived bias	0.16	-1.5	0	-4.3
Heterogeneous elasticity	0.16	$N(-1.5, 0.2)$	0	-4.3
2-product monopolist	0.16	-1.5 $(\epsilon_{ij} = 0.1)$	0	-3.0
Panel B: If the firm underestimates the bias ($\hat{\theta} < \theta$)				
Point estimates	0.16	-1.5	0.004	-2.7
Extreme estimates	0.10	-4.0	0.03	-1.1
Heterogeneous bias	$N(0.16, 0.05)$	-1.5	0.004	-3.2
Heterogeneous perceived bias	0.16	-1.5	$N(0.004, 0.01)$	-2.2
Heterogeneous elasticity	0.16	$N(-1.5, 0.2)$	0.004	-2.7
2-product monopolist	0.16	-1.5 $(\epsilon_{ij} = 0.1)$	0.004	-1.7

The table shows calculated percentage of gross profits lost, relative to optimal pricing. Optimal means that the firm perceives the bias to be what it actually is, $\hat{\theta} = \theta$, while the other cases take a lower perceived bias into account. Panel A shows the profits lost if the firm ignores the bias completely, and panel B shows the profits lost when it partially underestimate it. Each row makes a different assumption about the "true" parameters and the perceived ones (but always assuming that the firm knows the true price elasticity). For example, under the point estimates of left-digit bias of 0.16 but perceived bias of 0.004, and elasticity of -1.5, 2.7 percentage points are lost relative to optimal pricing. If the firm were to ignore the bias altogether, and assume it is non-existing, they would lose 4.3 percentage points.

The first row in each panel shows the lost profits under point estimates taken from the main analysis. The second row, extreme estimates, makes a firm-forthcoming assumption with lower bias, high price elasticity, and the upper bound of perceived bias. The next three rows explore the role of heterogeneity and how consequential it might be: I examine heterogeneity in left-digit bias, in perceived left-digit bias (truncating from below at 0.001), and in same price elasticity. Finally, the last row takes cross-elasticities into account (I use the implied perceived bias of 0.004, even though the implied perceived bias with cross-elasticities, given pricing, is lower. This way I am consistent and provide a lower bound on losses).

At each cell I simulate 3000 costs from a cost distribution that mimics the actual distribution inferred from a constant markup prices. For the heterogeneous cases I draw 50 values for each parameter. For the last row, I sample 80 prices for each of 2 products (from the same cost distribution as above).

A Proofs

A.1 Proof - Proposition 1

To see that equation 3 describes the optimal price, we will proceed in the following steps.

A.1.1 Interior solution

Consider the classic setting. There, since demand is convex, it satisfies the second order condition globally, and hence the first order condition is sufficient to characterize the solution. Here, the second order condition is satisfied locally, meaning that any interior solution is a local maximum. However, demand is not convex when changing the first digits. Therefore, one must also check for other prices, where convexity is not satisfied.

The first order condition is

$$\frac{\partial \Pi}{\partial p} = 0 = A\epsilon \left((1-\theta)\,p + \theta p_1\right)^{\epsilon-1}(1-\theta)\,(p-c) + A\left((1-\theta)\,p + \theta p_1\right)^{\epsilon}$$

$$0 = \epsilon\,(1-\theta)\,(p-c) + (1-\theta)\,p + \theta p_1$$

$$p = c\frac{\epsilon}{1+\epsilon} - p_1\frac{\theta}{1-\theta}\frac{1}{1+\epsilon}$$

$$p = \left(c - \frac{\theta}{1-\theta}\frac{q_1+1}{\epsilon}\right)\frac{\epsilon}{1+\epsilon}$$

Where the last transition is due to that $p_1 = q_1 + 1$ given a range of costs as in proposition 3, So, first, for any cost c, we can find a local maximum being the interior solution, as presented in the second row of equation 3. Note that this solution is strictly increasing in c, as long as p_1 does not change.

A.1.2 First case, $c < \bar{c}_q$

Since demand is locally convex, and isoprofits are strictly convex, it is sufficient to rule-out an interior solution as being optimal if the profits at the highest price below the discontinuity are higher. So, if $p^*(c)$ is the interior solution and $c < \bar{c}_q$, we need to show that $D(q)(q-c) > D(p^*)(p^*-c)$ where $q = p_1 - 0.01$:

$$D(p^*)(p^*-c) + D(p^*)(c-\bar{c}_q) = D(p^*)(p^*-\bar{c}_q)$$
$$< D(P)(P-\bar{c}_q)$$
$$= D(q)(q-\bar{c}_q)$$
$$= D(q)(q-c) + D(q)(c-\bar{c}_q)$$
$$\Leftrightarrow D(p^*)(p^*-c) < D(q)(q-c) + (D(q) - D(p^*))(c-\bar{c}_q)$$
$$< D(q)(q-c)$$

Where the first inequality (second row) is due to P being the optimal price for $\bar{c}_q$, and the last inequality (last row) is since $c < \bar{c}_q$ and $D(q) > D(p^*)$.

Once we have shown that q is more profitable than p^*, to see that the optimal price in that case is q itself, note that since demand is locally strictly convex between discontinuity points, if the profit at q is higher

than at the interior solution it means that the slope of the isoprofit (in the quantity-price dimension) is less steep than the demand curve, meaning that profits at q must be maximal for these costs. This is only true for θ low enough, such that there is not another point where demand is crossing the isoprofit[36].

A.1.3   Second case, $c > \bar{c}_q$

We want to show that if $c > \bar{c}_q$, then the interior solution is more profitable than q. That is, that $\Pi\left(p^{*}, c\right) > D\left(q\right)\left(q - c\right)$. Define p' as the price that given demand $D\left(q\right)$ results in the same profits as $\Pi^{*} = \Pi\left(p^{*}, c\right)$. That is,

$$
\begin{aligned}
p' &= \frac{\Pi^{*}}{D\left(q\right)} + c \\
&= \frac{\Pi^{*} - \Pi_{q}}{D\left(q\right)} + q + \left(c - \bar{c}_q\right) \\
\Rightarrow p' - q &= \frac{\Pi^{*} - \Pi_{q}}{D\left(q\right)} + \left(c - \bar{c}_q\right) \\
&> \left(c - \bar{c}_q\right)\left(1 - \frac{D\left(P\right)}{D\left(q\right)}\right) > 0
\end{aligned}
$$

where we used that $\Pi^{*} = \Pi\left(p^{*}, c\right) > \Pi\left(P, c\right) = \Pi_{q} - D\left(P\right)\left(c - \bar{c}_q\right)$, and that $D\left(P\right) < D\left(q\right)$. So, since $p' > q$ it means that $\Pi\left(p^{*}, c\right) = D\left(q\right)\left(p' - c\right) > D\left(q\right)\left(q - c\right)$.

A.2   Proof - Proposition 2

Rewriting equation 5,

$$
\begin{aligned}
P + \frac{(1-\theta)P + \theta P_{1}}{\epsilon\left(1-\theta\right)} - \frac{\left((1-\theta)P + \theta P_{1}\right)^{\epsilon} P - \left((1-\theta)q + \theta q_{1}\right)^{\epsilon} q}{\left((1-\theta)P + \theta P_{1}\right)^{\epsilon} - \left((1-\theta)q + \theta q_{1}\right)^{\epsilon}} &= 0 \\
P\left(1 + \frac{1}{\epsilon}\right) + \frac{T P_{1}}{\epsilon} - \frac{\left(P + T P_{1}\right)^{\epsilon} P - \left(q + T q_{1}\right)^{\epsilon} q}{\left(P + T P_{1}\right)^{\epsilon} - \left(q + T q_{1}\right)^{\epsilon}} &= 0 \\
P\left(1 + \frac{1}{\epsilon}\right) + \frac{T P_{1}}{\epsilon} - \frac{\Lambda^{\epsilon} P - q}{\Lambda^{\epsilon} - 1} &= 0 \\
P\left(1 + \epsilon\right)\left(\Lambda^{\epsilon} - 1\right) + T P_{1}\Lambda^{\epsilon} - T P_{1} &= \Lambda^{\epsilon} P \epsilon - q \epsilon \\
\left(P + T P_{1}\right)\Lambda^{\epsilon} &= P + T P_{1} + \left(P - q\right)\epsilon \\
\Lambda^{\epsilon} &= 1 + \frac{\left(P - q\right)\epsilon}{P + T P_{1}} \qquad (10)
\end{aligned}
$$

where $T \equiv \frac{\theta}{1-\theta}$, the likelihood ratio of being inattentive and $\Lambda \equiv \frac{P + T P_{1}}{q + T q_{1}}$.

---

[36] i.e., θ needs to satisfy that $\Pi\left(q, \underline{c}_q\right) \geq \Pi\left(q - 1, \underline{c}_q\right)$

A.2.1 The effect of elasticity on nl-price

First, note that $\frac{\partial \Lambda}{\partial \epsilon} = \frac{\partial P}{\partial \epsilon} \frac{1}{q + Tq1}$. Then, differentiating equation 10 with respect to ϵ,

$$(P + TP_1) \Lambda^\epsilon \left(\frac{\epsilon \frac{\partial P}{\partial \epsilon}}{P + TP_1} + \ln \Lambda \right) + \frac{\partial P}{\partial \epsilon} \Lambda^\epsilon = \frac{\partial P}{\partial \epsilon} + \frac{\partial P}{\partial \epsilon} \epsilon + P - q$$

$$\frac{\partial P}{\partial \epsilon} (\Lambda^\epsilon - 1)(1 + \epsilon) + (P + TP_1) \Lambda^\epsilon \ln \Lambda = P - q$$

$$\frac{\partial P}{\partial \epsilon} = \frac{P - q - (P + TP_1) \Lambda^\epsilon \ln \Lambda}{(\Lambda^\epsilon - 1)(1 + \epsilon)} > 0$$

To see that note that the denominator is positive for $\epsilon < -1$, since $\Lambda > 1$. Further, the enumerator is positive iff

$$(P - q)(q + Tq_1)^\epsilon - (P + TP_1)^{\epsilon+1} \ln \left(\frac{P + TP_1}{q + Tq_1} \right) > 0$$

It is easy to see that the above term equals zero for $P = q$. We now show that it is increasing in P, and hence positive, since $P > q$. To see that it is increasing in P, differentiate to show

$$(q + Tq_1)^\epsilon - (\epsilon + 1)(P + TP_1)^\epsilon \ln \left(\frac{P + TP_1}{q + Tq_1} \right) - (P + TP_1)^\epsilon > 0$$

$$\Leftrightarrow 1 > \Lambda^\epsilon (1 + (1 + \epsilon) \ln \Lambda)$$

and note that $\Lambda^\epsilon < 1$, and that $\ln \Lambda > 0$ and $\epsilon < -1$ hence $(1 + (1 + \epsilon) \ln \Lambda) < 1$.

A.2.2 Proof of corollary, effect of dollar digit on nl-price

We know that the NLP satisfies:

$$((1 - \theta)P + \theta P_1)^\epsilon \left(-\frac{(1 - \theta)P + \theta P_1}{\epsilon(1 - \theta)} \right) - ((1 - \theta)q + \theta q_1)^\epsilon \left(q - P \cdot \frac{1 + \epsilon}{\epsilon} - \frac{\theta P_1}{(1 - \theta)\epsilon} \right) = 0 \qquad (11)$$

$$\hat{p}^\epsilon \frac{\hat{p}}{\epsilon(1 - \theta)} + \hat{q}^\epsilon \left(q - p - \frac{\hat{p}}{\epsilon(1 - \theta)} \right) = 0 \qquad (12)$$

$$\hat{q}^\epsilon (q - p) \epsilon (1 - \theta) = \hat{p}(\hat{q}^\epsilon - \hat{p}^\epsilon) \qquad (13)$$

Where the transitions are just algebraic simplifications, and writing the perceived price instead of its components. To prove the corollary I ask if it is possible that the NLP will be that same for two consecutive dollar digits, $p_{q+1} = p_q + 1$. That is if,

$$(\hat{q} + 1)^\epsilon [(q - p) \epsilon (1 - \theta)] \lesseqgtr (\hat{p} + 1) \left((\hat{q} + 1)^\epsilon - (\hat{p} + 1)^\epsilon \right)$$

$$(\hat{q} + 1)^\epsilon \left[\hat{p}(\hat{q}^\epsilon - \hat{p}^\epsilon) \hat{q}^{-\epsilon} \right] \lesseqgtr (\hat{p} + 1) \left((\hat{q} + 1)^\epsilon - (\hat{p} + 1)^\epsilon \right)$$

$$\frac{\hat{p}}{\hat{p} + 1} \lesseqgtr \frac{1 - \left(\frac{\hat{p}+1}{\hat{q}+1} \right)^\epsilon}{1 - \left(\frac{\hat{p}}{\hat{q}} \right)^\epsilon}$$

if $\epsilon = -1$ it holds with equality:

$$\frac{\hat{p}}{\hat{p}+1} = \frac{\frac{\hat{p}+1-\hat{q}-1}{\hat{p}+1}}{\frac{\hat{p}-\hat{q}}{\hat{p}}} = \frac{\hat{p}}{\hat{p}+1}$$

but if $\epsilon < -1$, as assumed all over, then the RHS is lower. Meaning, that the equality does not hold.

A.2.3 The effect of inattention on nl-price

How does higher inattention affects the next lowest price? Differentiate equation 10 with respect to θ:

$$(\epsilon+1)\left(\frac{P+TP_1}{q+Tq_1}\right)^\epsilon \frac{\left(\frac{\partial P}{\partial\theta}+\frac{\partial T}{\partial\theta}P_1\right)(q+Tq_1)-\frac{\partial T}{\partial\theta}(P+TP_1)}{q+Tq_1} + \frac{\partial T}{\partial\theta}q_1\left(\frac{P+TP_1}{q+Tq_1}\right)^{\epsilon+1} - \frac{\partial P}{\partial\theta}(1+\epsilon) = \frac{\partial T}{\partial\theta}P_1$$

$$(\epsilon+1)\left(\frac{\partial P}{\partial\theta}+\frac{\partial T}{\partial\theta}P_1\right)+\frac{\partial T}{\partial\theta}(q_1-1-\epsilon)\frac{P+TP_1}{q+Tq_1} - \frac{\partial P}{\partial\theta}(1+\epsilon)\left(\frac{P+TP_1}{q+Tq_1}\right)^{-\epsilon} = \frac{\partial T}{\partial\theta}P_1\left(\frac{P+TP_1}{q+Tq_1}\right)^{-\epsilon}$$

$$\frac{\partial P}{\partial\theta}\left(1-\left(\frac{q+Tq_1}{P+TP_1}\right)^\epsilon\right) = \frac{1}{(1-\theta)^2}\left(\frac{P_1}{1+\epsilon}\left(\frac{q+Tq_1}{P+TP_1}\right)^\epsilon - P_1 + \left(1-\frac{q_1}{\epsilon+1}\right)\frac{P+TP_1}{q+Tq_1}\right)$$

the term in parentheses on the left-hand side is negative, so the effect is determined by the opposite sign of $\frac{P_1}{1+\epsilon}\left(\frac{q+Tq_1}{P+TP_1}\right)^\epsilon - P_1 + \left(1-\frac{q_1}{\epsilon+1}\right)\frac{P+TP_1}{q+Tq_1}$. This is sign is always negative for $\epsilon < -1$ for $q_1 > 0$.

To see that $\frac{P_1}{1+\epsilon}\left(\frac{q+Tq_1}{P+TP_1}\right)^\epsilon - P_1 + \left(1-\frac{q_1}{\epsilon+1}\right)\frac{P+TP_1}{q+Tq_1} < 0$, recall that $P - q = \Delta < 1$ and $P_1 = 1 + q_1$. Then, we need to show that

$$\frac{P_1}{1+\epsilon}\left(\frac{q+Tq_1}{P+TP_1}\right)^\epsilon + \left(1-\frac{q_1}{\epsilon+1}\right)\frac{P+TP_1}{q+Tq_1} < P_1$$

$$\frac{P+TP_1}{q+Tq_1}\left(\frac{1+\epsilon-q_1+P_1\left(\frac{q+Tq_1}{P+TP_1}\right)^{-(1+\epsilon)}}{1+\epsilon}\right) < 1+q_1$$

focusing on the LHS, using that $\epsilon < -1$ and $P_1 = 1 + q_1$:

$$\frac{P+TP_1}{q+Tq_1}\left(\frac{1+\epsilon-q_1+P_1\left(\frac{q+Tq_1}{P+TP_1}\right)^{-(1+\epsilon)}}{1+\epsilon}\right) < \frac{P+TP_1}{q+Tq_1}\left(\frac{2+\epsilon}{1+\epsilon}\right) < \frac{P+TP_1}{q+Tq_1}$$

then it is left to show that

$$\frac{P+TP_1}{q+Tq_1} = \frac{q+\Delta+T(1+q_1)}{q+Tq_1} < 1+q_1$$

or

$$q+\Delta+T(1+q_1) < (1+q_1)(q+Tq_1)$$
$$\Delta+T < q_1(q+Tq_1)$$

and since $\Delta + T \leq 1 + T$, if $q_1 \geq 1$, then $q + Tq_1 > 1 + T$. If $q_1 = 0$ then this inequality does not hold.

42

B Price Competition

In that section we solve for a price competition model where some consumers are inattentive to the first digit. However, the solution is for a continuous pricing, and the proof relies on the continuity of the price domain.

Varian (1980), with a different motivation and interpretation, solved a closely related model in his seminal paper "A Model of Sales". We solve a Varian-variant model, where there are N firms that have no costs of producing a homogeneous good. Prices can be set *continuously* between 0 and 2. A measure $1 - \theta$ of consumers is attentive, and a measure θ is partially inattentive. An attentive consumer buys from the cheapest firm (or randomizes between all cheapest firms.) A partially inattentive consumer has a coarse perception of the price, seeing the price as being "weakly less than 1" or "strictly more than 1". So a partially inattentive consumer is equally likely to purchase from any firm who prices below 1, or if no firm prices below 1, from those who price above 1.

Proposition 3. *There is no pure strategy equilibrium in this game*

Proof. Assume all firms play a strategy p. If $p > 0$ an ϵ downward deviation from a positive price will lend a firm the entire attentive market, while a price strategy of 0 is sub-optimal since setting the price at any $0 < p \leq 1$ gives positive profits from the partially inattentive market. □

Think of a mixed strategy as a cumulative distribution function F (with the accompanying PDF f) over prices and consider symmetric equilibria. Given F, notate $\phi \equiv F(1)$ as the probability of a firm pricing below 1. Define a *simple mixed equilibrium* as a symmetric equilibrium where $\phi = 1$, and a *truly mixed equilibrium* as a symmetric equilibrium where $\phi \in (0, 1)$.

Proposition 4. *(1) There exists a* simple mixed equilibrium *where $F(1) = 1$ and firms mix over $[p_0, 1]$ according to a unique $F_{\theta,N}$, $p_0 > 0$.*

(2) There sometimes exist a truly mixed equilibrium *as a function of θ and N, where firms mix over $[p_0, 1] \cup [p_1, 2]$ according to a unique $F_{\theta,N}$, where $p_0 > 0$ and $p_1 > 1$.*

An immediate corollary is that in any equilibrium, firms will not price at the point of discontinuity. Taking this simplistic, continuous domain model to our setting, the price of "1" represents the highest price before the drop in demand, i.e. prices that end with 99 before the policy or 90 post policy. So, that model predicts no pricing at 00.

Proof. Given ϕ from each firm's perspective, the number of other firms pricing under 1 is described by a binomial distribution $x \sim B(N - 1, \phi)$. It must be that $\phi > 0$, since if all firms price above 1, a deviation to 1 gives an expected payoff of 1, while the expected payoff from F must be strictly less than $\frac{2}{N}$ (since all firms have the same expected profit, and the maximal industry profit is 2). A firm's expected market share from a price $p \leq 1$ is

$$X(p|p \leq 1) = (1 - \theta)(1 - F(p))^{N-1} + \theta \mathbb{E}\left[\frac{1}{1 + x}\right]$$

$$= (1 - \theta)(1 - F(p))^{N-1} + \theta \frac{1 - (1 - \phi)^N}{N\phi}$$

A firm's expected market share from a price $p > 1$ is

$$X(p|p > 1) = (1 - \theta)(1 - F(p))^{N-1} + \theta \frac{(1 - \phi)^{N-1}}{N}$$

□

In each region, above or below 1, the CDF has no mass points. To see that, assume there is a mass point then it is profitable for a firm to shift that mass a little bit to the left[37].

Also, within each region there are no gaps. Assume there is, i.e. $f(p) = 0$ for all $p \in (\underline{p}, \overline{p})$ where either $\overline{p} < 1$ or $\underline{p} > 1$. Then, it is profitable to shift a probability from $f(p) > 0$ to a price in the gap region, thus only increasing profit without changing the probability of winning it. For the same reason $f(1) > 0$ and $f(2) > 0$.

Call an equilibrium where $\phi < 1$ a "truly mixed equilibrium". From the above, the support of the CDF in a truly mixed equilibrium is $[p_0, 1) \cup [p_1, 2]$. Now, we can find the CDF using the indifference principle, stating that the expected profit of every price point should be equal. Explicitly, it should be that the payoff is the same as for $p = 1$. So, the expected payoff is

$$(1 - \theta)(1 - \phi)^{N-1} + \theta \frac{1 - (1 - \phi)^N}{N\phi}$$

From that we can also find the lowest price p_0, since $F(p_0) = 0$. So,

$$p_0 \left((1 - \theta) + \theta \frac{1 - (1 - \phi)^N}{N\phi} \right) = (1 - \theta)(1 - \phi)^{N-1} + \theta \frac{1 - (1 - \phi)^N}{N\phi}$$

$$p_0 = \frac{(1 - \theta)(1 - \phi)^{N-1} + \theta \frac{1-(1-\phi)^N}{N\phi}}{(1 - \theta) + \theta \frac{1-(1-\phi)^N}{N\phi}}$$

Notice that if $\phi = 1$, all firms never price above 1, then $p_0 = \frac{\theta}{N(1-\theta)+\theta}$, as in Varian 1980.

What is p_1? $F(p_1) = \phi$, the price is higher than 1, but the expected share of inattentives is conditional on all firms pricing above 1. That is,

$$p_1 \left((1 - \theta)(1 - \phi)^{N-1} + \theta \frac{(1 - \phi)^{N-1}}{N} \right) = (1 - \theta)(1 - \phi)^{N-1} + \theta \frac{1 - (1 - \phi)^N}{N\phi}$$

$$p_1 = \frac{(1 - \theta)(1 - \phi)^{N-1} + \theta \frac{1-(1-\phi)^N}{N\phi}}{(1 - \theta)(1 - \phi)^{N-1} + \theta \frac{(1-\phi)^{N-1}}{N}}$$

---

[37]If there is a mass point at 1, then it is profitable to shift it to just below 1 and gain $\theta \frac{1-(1-\phi)^N - (1-\phi)^{N-1}\phi}{N\phi} = \theta \frac{1-(1-\phi)^{N-1}}{N\phi} > 0$

Next, notice that $F(2) = 1$, and from the indifference principle,

$$2\theta \frac{(1-\phi)^{N-1}}{N} = (1-\theta)(1-\phi)^{N-1} + \theta \frac{1-(1-\phi)^N}{N\phi}$$

$$(1-\phi)^{N-1}\left(2 - N\frac{1-\theta}{\theta} + \frac{1-\phi}{\phi}\right) = \frac{1}{\phi}$$

$$(1-\phi)^{N-1}\left(\phi\left(2 - N\frac{1-\theta}{\theta}\right) + (1-\phi)\right) = 1$$

This equation determines ϕ, if it even exists. Note that the leftmost term is smaller than 1, and the term in parentheses is a weighted average of 1 and $2 - N\frac{1-\theta}{\theta}$. If that last term is less than 1, then no ϕ can solve that equation. So, a necessary condition for such an equilibrium to exist is that

$$2 - N\frac{1-\theta}{\theta} > 1$$

$$N\frac{1-\theta}{\theta} < 1$$

$$N < \frac{\theta}{1-\theta} \qquad\qquad \theta > \frac{N}{1+N}$$

For example, if $N = 2$ then $\theta > \frac{2}{3}$. In practice, we assume that the share of inattentives is likely to be small. However, this restriction relies on the marginal costs being 0. A higher marginal cost $0 < c < 1$ will alter the above indifference,

$$(2-c)\theta\frac{(1-\phi)^{N-1}}{N} = (1-c)\left((1-\theta)(1-\phi)^{N-1} + \theta\frac{1-(1-\phi)^N}{N\phi}\right)$$

$$(1-\phi)^{N-1}\left(\phi\left(\frac{2-c}{1-c} - N\frac{1-\theta}{\theta}\right) + (1-\phi)\right) = 1$$

Therefore, the above constraint is relaxed to be $\theta > \frac{N}{\frac{1}{1-c}+N}$. For example, if $N = 2$ and $c = 0.5$, then $\theta > \frac{1}{2}$ instead of $\frac{2}{3}$. Still, we need a small N and high c to bring the bound on θ to reasonable values.

In any case, if $\phi = 1$ we are almost back to the Varian case, and an equilibrium exists. Call it a "simple mixed equilibrium". To see that, assume that all firms play F such that $F(1) = 1$. Then, deviating any density to above 1 gains nothing - both attentives and inattentives will not buy at that price. Therefore a simple mixed equilibrium, as already characterized by Varian, exists. This is actually a way to make the arbitrary assumption that Varian makes, of a bound on the maximal price, to be weakly less arbitrary.

C   Multi-product Monopolist

Even though single-item sellers use 9-ending prices as well (firms such as Netflix, or Amazon, or in house listings Repetto and Solis (2018)), supermarkets are sellers of multiple items. Substitution between differentiated products is a main focus and an empirical challenge that spans a huge literature (see review by Nevo (2011)). How does simultaneous pricing of multiple products with cross-elasticities between them affect the firm pricing behavior regarding 9-ending pricing?

To keep the problem simple, consider a firm that sells two products setting prices to maximize profit of the form

$$\Pi = D\left(p_i, p_j; \theta\right)(p - c)$$

where $D = (D_i, D_j)$ is a vector of demand for items i and j, $p = (p_i, p_j)$ is the prices vector, and $c = (c_i, c_j)$ is the unit costs vector. Let

$$D_i = A_i \hat{p}_i^{\epsilon_{ii}} \hat{p}_j^{\epsilon_{ij}}$$

where ϵ_{ii} is the same price elasticity, and $\epsilon_{ij} = \epsilon_{ji}$ is the cross-product elasticity. Then the firm's problem is to maximize

$$\Pi = \max_{p_i, p_j} A_i \hat{p}_i^{\epsilon_{ii}} \hat{p}_j^{\epsilon_{ij}}(p_i - c_i) + A_j \hat{p}_i^{\epsilon_{ij}} \hat{p}_j^{\epsilon_{jj}}(p_j - c_j)$$

If the cross-elasticity is 0, we are back to the single item monopolist, since the problem is separable. Deriving a closed-form solution is possible, but it is not very insightful, in the sense that I do not have any comparative statics to provide. Instead, I solve the problem numerically, and focus attention on the effects of adding cross elasticities on the region of missing prices (numerical, local, comparative statics). Figure A.10 provides an example for such effects. Here, the black curves describe the price versus cost of product i, when product j has a cost of 1.5, and the orange curves are the price of product j. Both items have same price elasticity of -2, and cross-elasticity is either zero (the solid lines) or 0.1 (the dashed line). Margins of product i are higher due to the substitution with product j. Also, the arrows point to the lowest-prices of 4.67 in the single item case, and 4.73 with the cross-elasticity.

The finding of a higher next-lowest price for higher substitution holds in more general simulations – when I allow for varying costs for the other item as well (taking the next-lowest price regardless of the other items underlying cost), different same-price elasticities for each product, and at various levels of left-digit bias. Further, it holds for both the more expensive and cheaper product. While this is hardly a proof, it seems to be a plausible conjecture.

D   Nielsen Data

Here I describe the Nielsen data, and discuss its issues and the solution I propose.

Nielsen Retail Scanner (RMS) and Consumer Panel (HMS) data are provided by the Kilts Center at the University of Chicago. RMS records weekly UPC-store level quantity and revenues. Summary statistics are described in column (1) of Appendix Table 1. The sample consists of 28 million observations of 12 products from DellaVigna and Gentzkow (2017), selling in 11000 stores, with a total of 188 million units sold for $881 million. The average price is at $5.51, with a standard deviation of 2.81.

The data is similar to the retailer's data, with the key advantage being that it includes the clientele and pricing behavior of more than 60 retailers rather than one. Another pragmatic advantage is the longer panel of 9 years instead of 3.5. However, the data have three main disadvantages, and as such is not part of the main analysis. First, prices are rounded such that they are always "to the cent", thus not allowing to infer if an observation represents a fictitious price. Second, Nielsen measurement weeks start on Sunday, and are not always aligned with a retailer's price changing frequency. For example, some chains change their prices weekly but do so on Wednesdays, so if a price changed the Nielsen price will mix sales with one price for Sunday-Tuesday and another for Wednesday-Saturday. Measurement misalignment leads to far more fictitious prices than in the retailer data. Finally, only the weekly average price is observed, excluding

a simple way to infer if a price is on-sale or not. Additional variables, "promotion" and "display", are available but are missing for 80% of the observations.

To solve the issues of fictitious prices and sale prices I use a new, yet simple, data cleaning technique. If an observed price of a unique product in a certain store is the result of averaging different prices, it is unlikely to be identical to the price in the following week [38]. For example, imagine there is a discount to club members only, so some consumers pay $4.99 and others $4.29. In order for the "weekly" price to be the same across weeks, the share of club members among weekly consumers must be kept exactly the same. Therefore, I define for each product-store the spell length of each price and exclude 1-week spells. Longer spells are also more likely to be regular prices, or conversely, shorter spells are more likely to be on-sale prices. Therefore, while I can not directly observe if a product is on-sale in a certain week, I proxy for it with the spell length. The spell length proxy has the advantage of not inferring an on-sale by price, which is problematic since regular prices also change.

These intuitive arguments are corroborated using the retailer data, where they can be directly examined. In Appendix Figure A.3 I am testing these conjectures for the 1710 products of the retailer data. The black circles show the share of observations that are real (to the cent) conditional on the spell length. About 50% are real for a 1 week spell, and that share jumps to 95% for 2 week spells, and gets close to a 100% for 3 weeks and longer. The yellow triangles show that share of observations that are regular prices (not on-sale) by spell length. About 25% of 1-week spells represent regular prices, 40% at 2 weeks, and by 6 week spells almost all are regular prices.

The price ending distributions, depicted in Figure A.1, further depicts and support the data cleaning procedure. To recap, non-real prices are excluded in the retailer data by excluding non-round prices, and in Nielsen data by excluding 1 week spells. An observation is on-sale in the retailer data if the net price is lower than the regular price, and in Nielsen I add dummies for 2,3,4, and 5 week spells to proxy for on-sale propensity (motivated by Appendix Figure A.3). Indeed, a similar qualitative pattern appears in both data sets even though the data cleaning procedure is different. The figure shows the shares of observations that end with 9, 0, or other price-ending, and each color bars describe a different cut of the data. Moving from all observations to a sample that includes only "real" prices (to-the-cent or at least 2 weeks spells), decreases the shares of non-9 non-0 observations, which add almost solely to 9-ending prices. i.e., without that exclusion, about a third to half of observations in other price endings should be attributed to 9-ending prices. Further focusing on likely regular prices (observable in the retailer data, and defined as at least 6 week spells), increases the share of 9-ending prices, but on the expense of 0-ending ones. That is, 0-ending prices are mostly on-sale prices or more transient price spells.

E   Bias Heterogeneity

Armed with estimates of the bias at the product level, we can ask how does the bias change according to the product characteristics? A few mechanisms are interesting to look at. First, a mechanism of relative thinking might be in place. The idea was introduced by Tversky and Kahneman (1981) with the famous calculator-jacket example. Namely, 68% of respondents take the same hypothetical effort for a $5 discount if it is out of a $15 item, while only 29% if it is out of a $125 item. Needless to say, this kind of behavior is not consistent with utility maximization. While different theories were developed (e.g., Azar (2007); Bushong et al. (2015)), field evidence for the effect is rare (with the recent exception of Hirshman et al. (forthcoming)).

---

[38]Mostly true for a popular product purchased in multiple units each week, otherwise might be likely.

In this left-digit bias setting, a manifestation of relative thinking is if the bias is higher for high ticket items, everything else constant. Meaning, if a cent out of a $2 item weighs more heavily than a cent out of an $8 item. Already hinted in Figure 3, the drops are not shrinking for higher dollar digits. According to the sufficient statistic formula, with the same elasticity and demand drop size, the bias is increasing linearly with the dollar digit. Thus, it seems likely to suspect that relative thinking might be affecting the bias.

Higher bias for more expensive products is irrational if a single unit is purchased of each item. However, if cheaper items are also purchased more frequently and in multiple units by purchasing households, the price level correlation might be a proxy for a form of rational inattention. Therefore, I turn to Nielsen HomeScan Panel (HMS), and calculate for each product the average number of units purchased of an item, conditional on it being purchased in a shopping trip[39]. If the bias is a form of deliberate inattention, focusing more on products whose last-digits weight on the budget is higher, we should expect more frequently purchased items to be correlated with lower bias.

Finally, one might be curious how the demographics of purchasing clientele are related to the bias. I focus on two main variables, household income and education, also taken as the average level of those by purchasing households in the HMS. Income has an ex-ante unknown sign, since low income might be correlated with more behavioral biases (CITES), but high income people might care less about the cents (recall, the bias is above and beyond the overall product price elasticity). Education is expected to be more clearly, correlated with less bias (e.g. Bronnenberg et al. (2015)).

Since I have various measures of the bias from the three methods discussed above, each with its benefits and drawbacks, I use a simple mean as the bias measure. This is a crude way of minimizing sampling error in the estimates. The simple correlates appear in Figure A.8. As hinted above, price (panel a), is positively correlated with the bias; demand is negatively correlated (panel b); and education and income are less clearly correlated (panels c and d).

More important than the univariate correlations, is how are these affected when taken together. In a multivariate case we can look at the price effects controlling for number of units purchased, or on the education effects controlling for income. Table A.2 show these results. Columns 1-4 are the univariate regressions of the bias on the four different variables, while columns 5-7 are multivariate versions taking all variables together. Column 7 adds subcategory fixed effects to look at the variation of the bias within products of the same type (e.g., "coffee", "rice", "cottage cheese", or "refrigerated orange juice"). The overall picture is that both price and number of units are associated with the bias, where a 1% higher price is associated with a 0.21 (standard error 0.04) points higher bias (that is additional 2.1 cents are ignored), while a 1% higher number of units purchased is associated with 0.22 (0.09) points lower bias. Income is positively correlated with the bias, and education negatively so, but the estimates are not significantly different than zero. Meaning, we find suggestive support for both relative thinking and deliberate inattention.

F   Appendix Figures

References

Aalto-Setala, Ville, "How do markets behave? The adjustment of price endings," *Journal of Product & Brand Management*, 2005, *14* (7), 455–459.

---

[39]I also calculate the total number of units purchased within a year, the results are very similar, but the per-trip purchases seem like a more relevant measure.
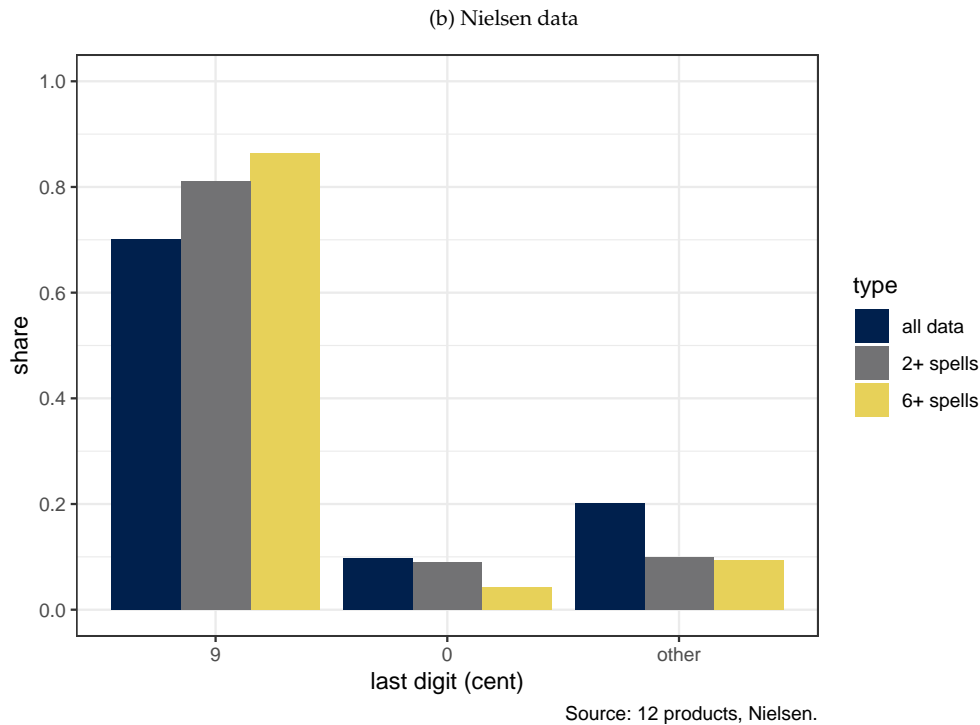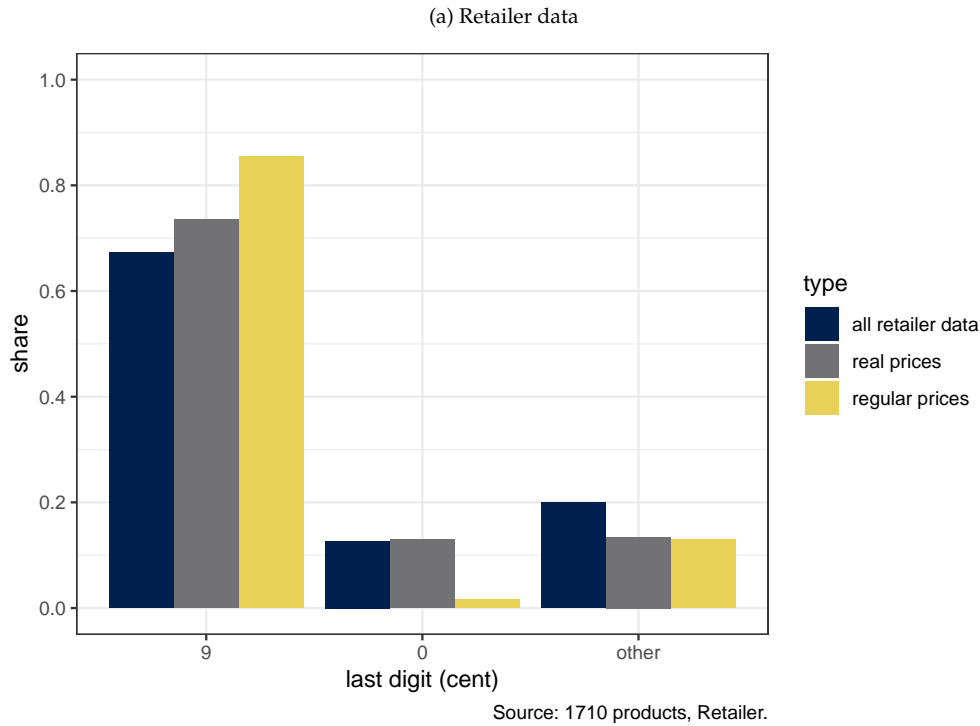
Akerlof, George A and Robert J Shiller, *Phishing for phools: The economics of manipulation and deception*, Princeton University Press, 2015.

Anderson, Eric, Nir Jaimovich, and Duncan Simester, "Price stickiness: Empirical evidence of the menu cost channel," *Review of Economics and Statistics*, 2015, *97* (4), 813–826.

Anderson, Eric T. and Duncan I. Simester, "Effects of $9 Price Endings on Retail Sales: Evidence from Field Experiments," *Quantitative Marketing and Economics*, 2003, *1* (1), 93–110.

Ashton, Lydia, "Left-digit bias and inattention in retail purchases: evidence from a field experiment," 2014. Available at SSRN: https://ssrn.com/abstract=2538816.

Ater, Itai and Omri Gerlitz, "Round prices and price rigidity: Evidence from outlawing odd prices," *Journal of Economic Behavior & Organization*, 2017, *144*, 188–203.

Azar, Ofer H, "Relative thinking theory," *The Journal of Socio-Economics*, 2007, *36* (1), 1–14.

Basu, Kaushik, "Why are so many goods priced to end in nine? And why this practice hurts the producers," *Economics Letters*, 1997, *54* (1), 41–44.

__ , "Consumer Cognition and Pricing in the Nines in Oligopolistic Markets," *Journal of Economics & Management Strategy*, 2006, *15* (1), 125–141.

Bizer, George Y. and Robert M. Schindler, "Direct Evidence of Ending-Digit Drop-Off in Price Information Processing," *Psychology and Marketing*, 2005, *22* (10), 771–783.

Blattberg, Robert C and Kenneth J Wisniewski, *Issues in Modeling Store Level Scanner Data*, University of Chicago, Center for Research in Marketing, 1988.

__ **and Scott A Neslin**, *Sales promotion: Concepts, methods, and strategies*, Prentice Hall, 1990.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

Bray, Jeffery Paul and Christine Harris, "The effect of 9-ending prices on retail sales: A quantitative UK based field study," *Journal of Marketing Management*, 2006, *22* (5-6), 601–617.

Bronnenberg, Bart J, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M Shapiro, "Do pharmacists buy Bayer? Informed shoppers and the brand premium," *The Quarterly Journal of Economics*, 2015, *130* (4), 1669–1726.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein, "A model of relative thinking," *Unpublished manuscript, Harvard University, Cambridge, MA*, 2015.

Busse, Meghan R, Nicola Lacetera, Devin G Pope, Jorge Silva-Risso, and Justin R Sydnor, "Estimating the effect of salience in wholesale and retail car markets," *American Economic Review*, 2013, *103* (3), 575–79.

Carver, James R and Daniel T Padgett, "Product category pricing and future price attractiveness: 99-ending pricing in a memory-based context," *Journal of Retailing*, 2012, *88* (4), 497–511.

Cavallo, Alberto, "Scraped data and sticky prices," *Review of Economics and Statistics*, 2016, (0).

__ **and Roberto Rigobon**, "The billion prices project: Using online prices for measurement and research," *Journal of Economic Perspectives*, 2016, *30* (2), 151–78.

Chava, Sudheer and Vincent W Yao, "Cognitive Reference Points, the Left-Digit Effect, and Clustering in Housing Markets," 2017.

Chen, Yuxin, Ganesh Iyer, and Amit Pazgal, "Limited memory, categorization, and competition," *Marketing Science*, 2010, *29* (4), 650–670.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

Cho, Sungjin and John Rust, "Is econometrics useful for private policy making? A case study of replacement policy at an auto rental company," *Journal of econometrics*, 2008, *145* (1-2), 243–257.

＿ **and** ＿ , "The flat rental puzzle," *The Review of Economic Studies*, 2010, *77* (2), 560–594.

Conlon, Christopher T and Nirupama Rao, "Discrete Prices and the Incidence and Efficiency of Excise Taxes," 2016.

Dalrymple, Douglas J and George H Haines Jr, "A study of the predictive ability of market period demand-supply relations for a firm selling fashion products," *Applied Economics*, 1970, *1* (4), 277–285.

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 2009, *47* (2), 315–372.

＿ , "Structural Behavioral Economics," in Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *The Handbook of Behavioral Economics*, Elsevier, forthcoming.

＿ **and Matthew Gentzkow**, "Uniform Pricing in US Retail Chains," 2017.

＿ **and Ulrike Malmendier**, "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

＿ **and** ＿ , "Paying Not to Go to the Gym," *American Economic Review*, 2006, *96* (3), 694–719.

Dube, Arindrajit, Alan Manning, and Suresh Naidu, "Monopsony and Employer Mis-optimization Account for Round Number Bunching in the Wage Distribution," *Unpublished manuscript*, 2017.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo, "Reference prices, costs, and nominal rigidities," *The American Economic Review*, 2011, *101* (1), 234–262.

＿ **,** ＿ **,** ＿ **, and Josephine Smith**, "How frequent are small price changes?," *American Economic Journal: Macroeconomics*, 2014, *6* (2), 137–155.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo, "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?," *QME*, 2010, *8* (2), 207–239.

Farhi, Emmanuel and Xavier Gabaix, "Optimal taxation with behavioral agents," Technical Report, National Bureau of Economic Research 2015.

Gabaix, Xavier, "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

Ginzberg, Eli, "Customary prices," *The American Economic Review*, 1936, *26* (2), 296–296.

Goldfarb, Avi and Mo Xiao, "Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets," *American Economic Review*, 2011, *101* (7), 3130–61.

Gopinath, Gita, Pierre-Olivier Gourinchas, Chang-Tai Hsieh, and Nicholas Li, "International prices, costs, and markup differences," *The American Economic Review*, 2011, *101* (6), 2450–2486.

Hackl, Franz, Michael E Kummer, and Rudolf Winter-Ebmer, "99 Cent: Price points in e-commerce," *Information Economics and Policy*, 2014, *26*, 12–27.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein, "Learning through noticing: Theory and evidence from a field experiment," *The Quarterly Journal of Economics*, 2014, *129* (3), 1311–1353.

Hausman, Jerry A, "Valuation of new goods under perfect and imperfect competition," in "The economics of new goods," University of Chicago Press, 1996, pp. 207–248.

Hausman, Jerry, Gregory Leonard, and J Douglas Zona, "Competitive analysis with differenciated products," *Annales d'Economie et de Statistique*, 1994, pp. 159–180.

Heidhues, Paul and Botond Kőszegi, "Behavioral Industrial Organization," in Douglas Bernheim, Stefano DellaVigna, and DavidEditors Laibson, eds., *The Handbook of Behavioral Economics*, Elsevier, forthcoming.

Hirshman, Samuel, Devin Pope, and Jihong Song, "Mental Budgeting vs. Relative Thinking," *American Economic Association Papers and Proceedings*, forthcoming.

Hossain, Tanjim and John Morgan, "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," *The B.E. Journal of Economic Analysis & Policy*, 2006, *5* (2), 1–30.

Kalyanam, Kirthi and Thomas S Shively, "Estimating irregular pricing effects: A stochastic spline regression approach," *Journal of Marketing Research*, 1998, pp. 16–29.

Kleven, Henrik J and Mazhar Waseem, "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan," *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor, "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 2012, *102* (5), 2206–2236.

Levy, Daniel, Dongwon Lee, Haipeng Chen, Robert J. Kaufmann, and Mark Bergen, "Price Points and Price Rigidity," *Review of Economics and Statistics*, 2011, *94* (4), 1417–1431.

Macé, Sandrine, "The impact and determinants of nine-ending pricing in grocery retailing," *Journal of Retailing*, 2012, *88* (1), 115–130.

Matějka, Filip, "Rationally inattentive seller: Sales and discrete pricing," *The Review of Economic Studies*, 2015, *83* (3), 1125–1155.

Montgomery, Alan L, "Creating micro-marketing pricing strategies using supermarket scanner data," *Marketing science*, 1997, *16* (4), 315–337.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

—, "Empirical models of consumer behavior," *Annual Review of Economics*, 2011, *3* (1), 51–75.

**Repetto, Luca and Alex Solis**, "The Price of Inattention: Evidence from the Swedish Housing Market," 2018.

Rossi, Peter E, "Even the rich can make themselves poor: A critical examination of IV methods in marketing applications," *Marketing Science*, 2014, *33* (5), 655–672.

__ **, Greg M Allenby, and Rob McCulloch**, *Bayesian statistics and marketing*, John Wiley & Sons, 2012.

Saez, Emmanuel, "Do taxpayers bunch at kink points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Schindler, Robert M and Alan R Wiman, "Effects of odd pricing on price recall," *Journal of Business Research*, 1989, *19* (3), 165–177.

__ **and Patrick N Kirby**, "Patterns of rightmost digits used in advertised prices: implications for nine-ending effects," *Journal of Consumer Research*, 1997, *24* (2), 192–201.

__ **and Rajesh Chandrashekaran**, "Influence of price endings on price recall: a by-digit analysis," *Journal of Product & Brand Management*, 2004, *13* (7), 514–524.

__ **and Thomas M Kibarian**, "Increased consumer sales response though use of 99-ending prices," *Journal of Retailing*, 1996, *72* (2), 187–199.

Sehity, Tarek El, Erik Hoelzl, and Erich Kirchler, "Price developments after a nominal shock: Benford's Law and psychological pricing after the euro introduction," *International Journal of Research in Marketing*, 2005, *22* (4), 471–480.

Shlain, Avner, "Left-Digit Bias and Firm Sophistication," *Unpublished manuscript*, 2018.

Snir, Avichai, Daniel Levy, and Haipeng Allan Chen, "End of 9-endings, price recall, and price perceptions," *Economics Letters*, 2017, *155*, 157–163.

Spiegler, Ran, *Bounded rationality and industrial organization*, OUP USA, 2011.

Stiving, Mark, "Price-Endings When Prices Signal Quality," *Management Science*, 2000, *46* (12), 1617–1629.

__ **and Russell S. Winer**, "An Empirical Analysis of Price Endings with Scanner Data," *Journal of Consumer Research*, 1997, *24* (1), 57–67.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," Working Paper 22545, National Bureau of Economic Research August 2016.

Thomas, Manoj and Vicki Morwitz, "Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition," *Journal of Consumer Research*, 2005, *32* (1), 54–64.

Tversky, Amos and Daniel Kahneman, "The framing of decisions and the psychology of choice," *Science*, 1981, *211* (4481), 453–458.

Varian, Hal R, "A model of sales," *The American Economic Review*, 1980, *70* (4), 651–659.

Wansink, Brian, Robert J Kent, and Stephen J Hoch, "An anchoring and adjustment model of purchase quantity decisions," *Journal of Marketing Research*, 1998, pp. 71–81.

Figure A.1: Data comparison, price ending shares, and sample selection

(a) Retailer data

Source: 1710 products, Retailer.

(b) Nielsen data

Source: 12 products, Nielsen.

The figures show the shares of 9-ending prices, 0-ending prices and others in the Retailer data (panel A) and Nielsen data (panel B), and how these change by sample selection. The dark purple bars are the data before cleaning (corresponding to column (2) in Table 1 and column (1) in A.1). The middle green bars are the shares of "real" prices in the data (columns (3) and (2)) defined as to-the-cent prices in the Retailer data and at least 2 week price spells in Nielsen. The light yellow bars are the shares for "regular" prices (columns (4) and (3)), defined as instances in which the net price equals the regular price in the Retailer data and price spells of at least 6 weeks in Nielsen.

Figure A.2: Data description: "real" and "on-sale" by last digit of observed price
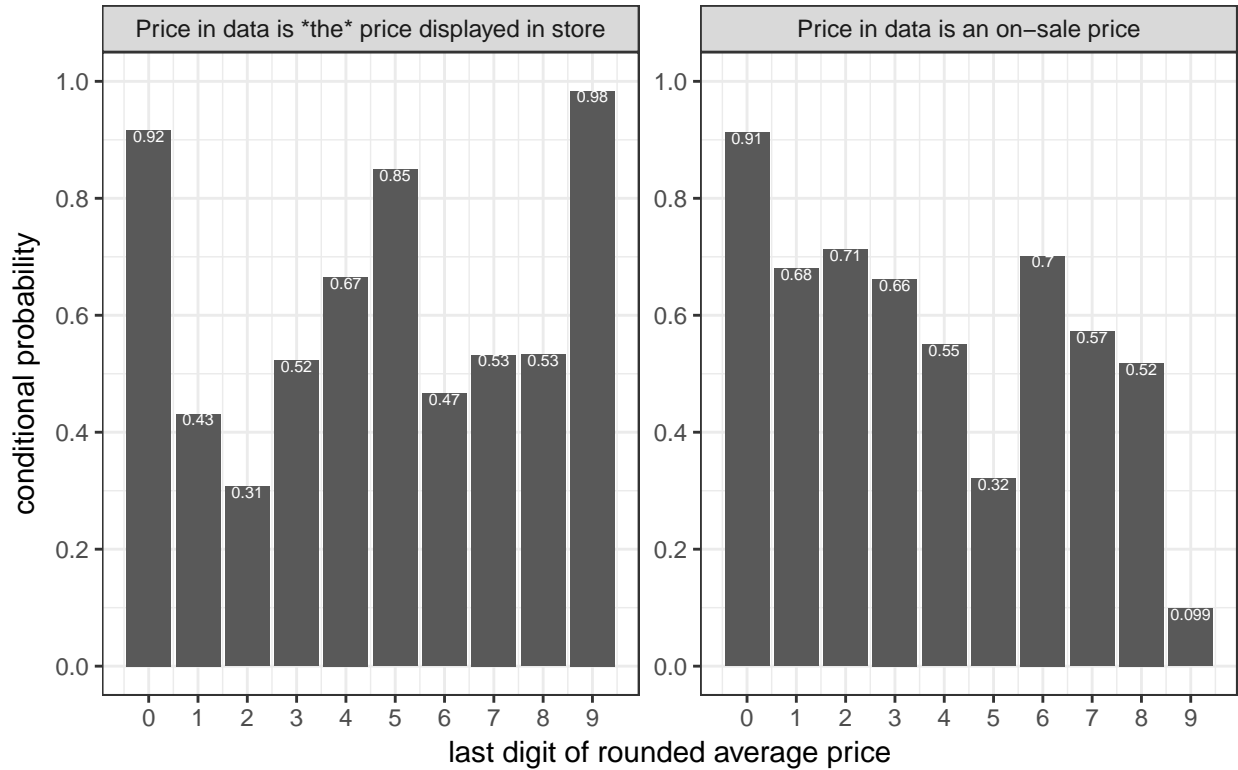
Based on 1710 products, 52M observations. `on–sale` = weekly average price is lower than regular price.
`real price` = the weekly average price is `to the cent`. Shares of price endings: 9: 67%, 0: 13%, 5: 6%, others <2%.
source: retailer data.

The figure shows the shares of observed prices that are to-the-cent (left) and on-sale (right), conditional on the last digit of the (rounded to the cent) price. For example, on the left, prices ending with 9 are 98% to-the-cent while 2-ending prices are to-the-cent in 31% of observations. Meaning, for 69% of observed 2-ending rounded price, the observation represents a mixture of several prices. On the right, an observation is on-sale if the actual price is lower than the regular price. A 9-ending price is on-sale 9.9% of the time, while a 0-ending price is on-sale for 91% of these observations.

Figure A.3: Shares of real prices and on-sale prices by length of price spell
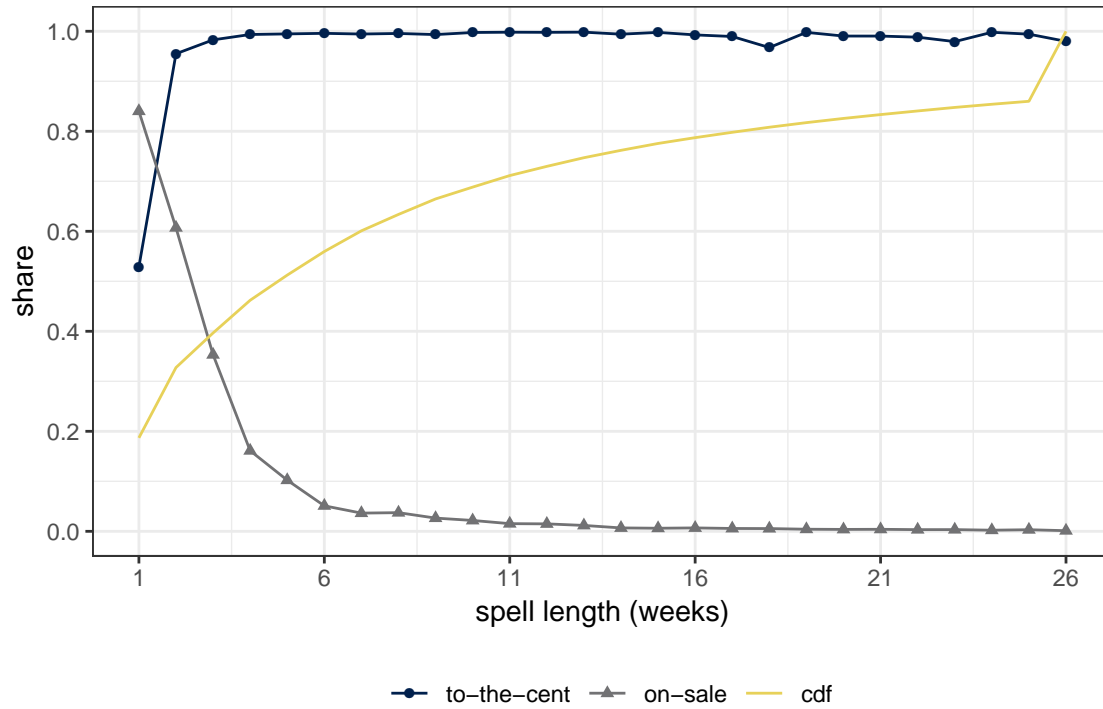
Shares of observations that are to-the-cent prices and on-sale by spell-length for the retailer data (corresponding to column (3) in Table 1). A product-week-store observation is allocated to spell length T if the price on that week is identical across T consecutive weeks. The light yellow line is the CDF of observations in the data. The dark circles are the share of observations, conditional on spell length, that are "real" (i.e., prices ending to-the-cent). The lighter triangles are the share of observations, conditional on spell length, that are on-sale (i.e. net price is less than the regular price).

Figure A.4: Retailer Robustness: Residualized demand of 12 products and 70 retailers using Nielsen data
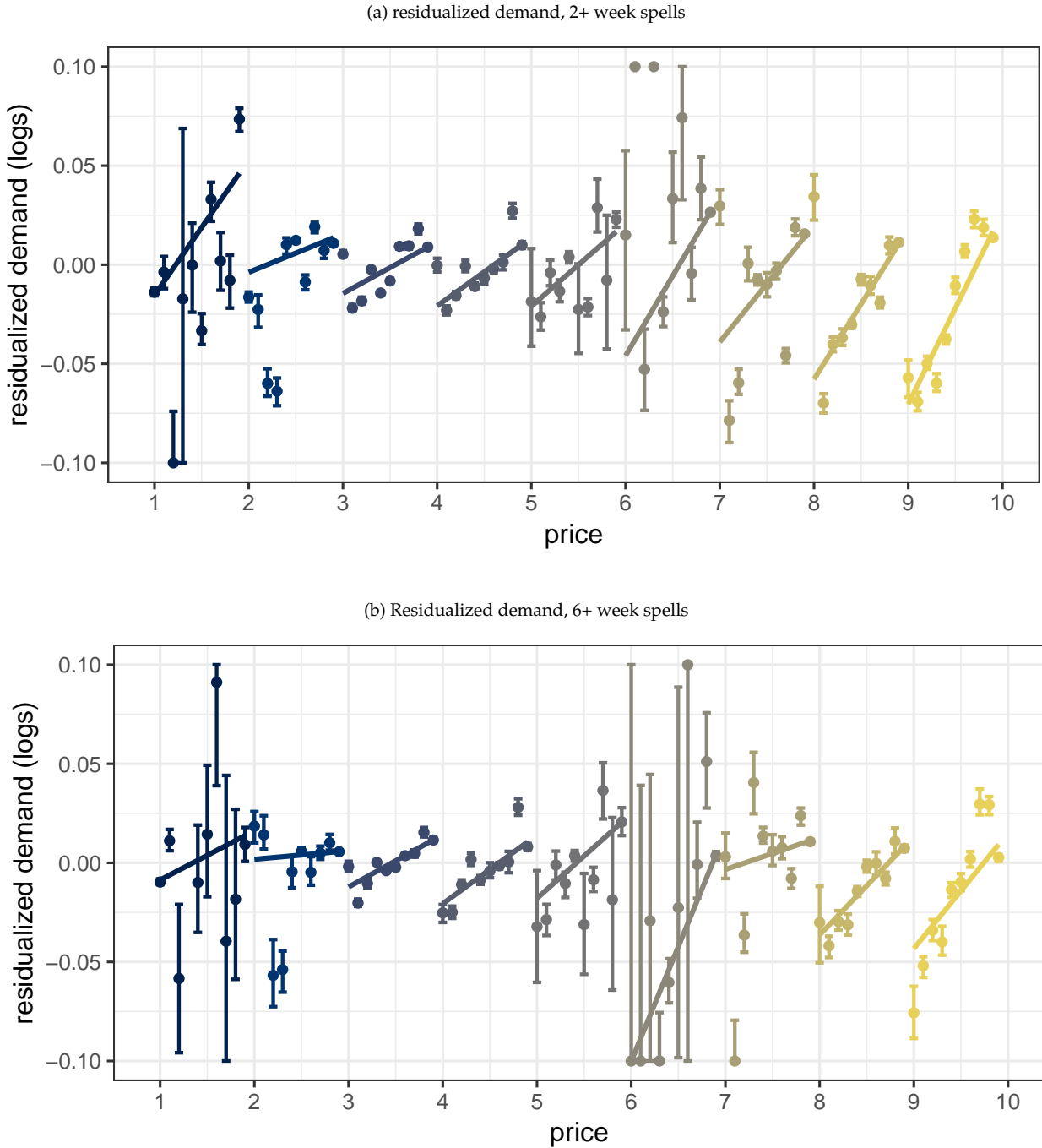
(a) residualized demand, 2+ week spells
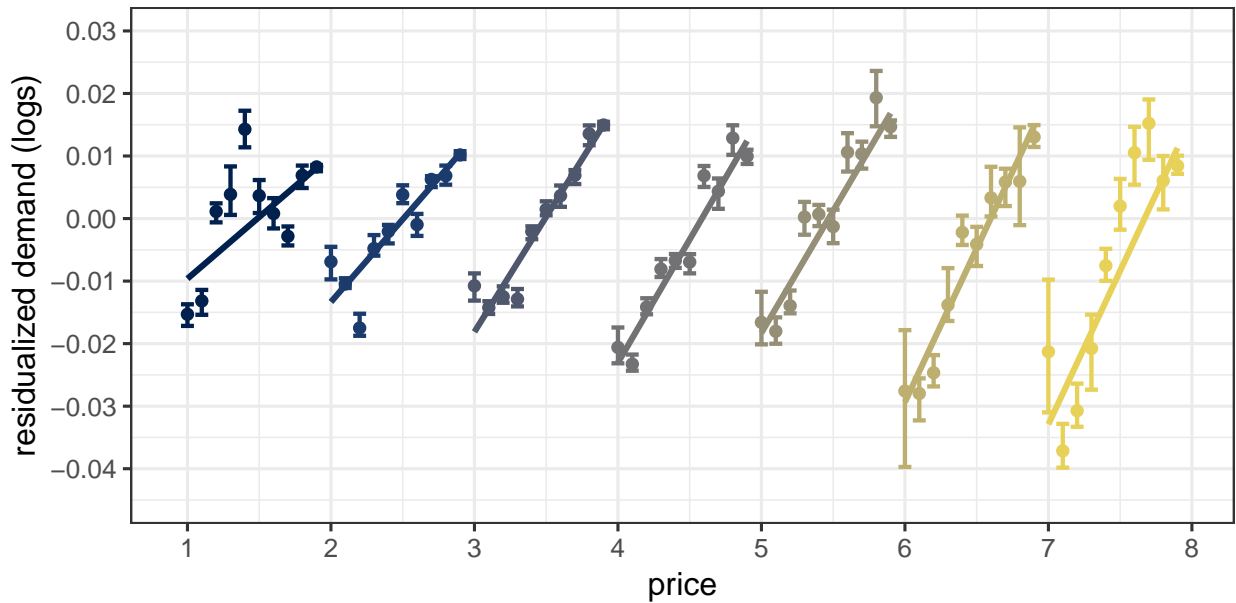

(b) Residualized demand, 6+ week spells

Net log-demand by price, residualizing product-level price elasticity, seasonality, store fixed effects, and promotion effects. Regressions are conducted at the product-chain level (12 products, 11000 stores of 70 retailers, 468 weeks). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Fixed effects are estimated at 10-cent bins, and 95% confidence intervals are result of 300 cluster-bootstraps by store.

The top panel is using data on all observations of at least 2-week spells (Column (2) in Table A.1). The bottom panel presents the results from a restricted sample of at least 6-week spells (column (3) of Table A.1). Data is truncated for prices between $1 and $9.99 for clarity. For further details see Section 4.1 for the empirical specification and Appendix D for data description.

Figure A.5: Demand structure robsutness: Residualized log-linear demand

(a) Main sample, constant semi-elasticity IV estimation
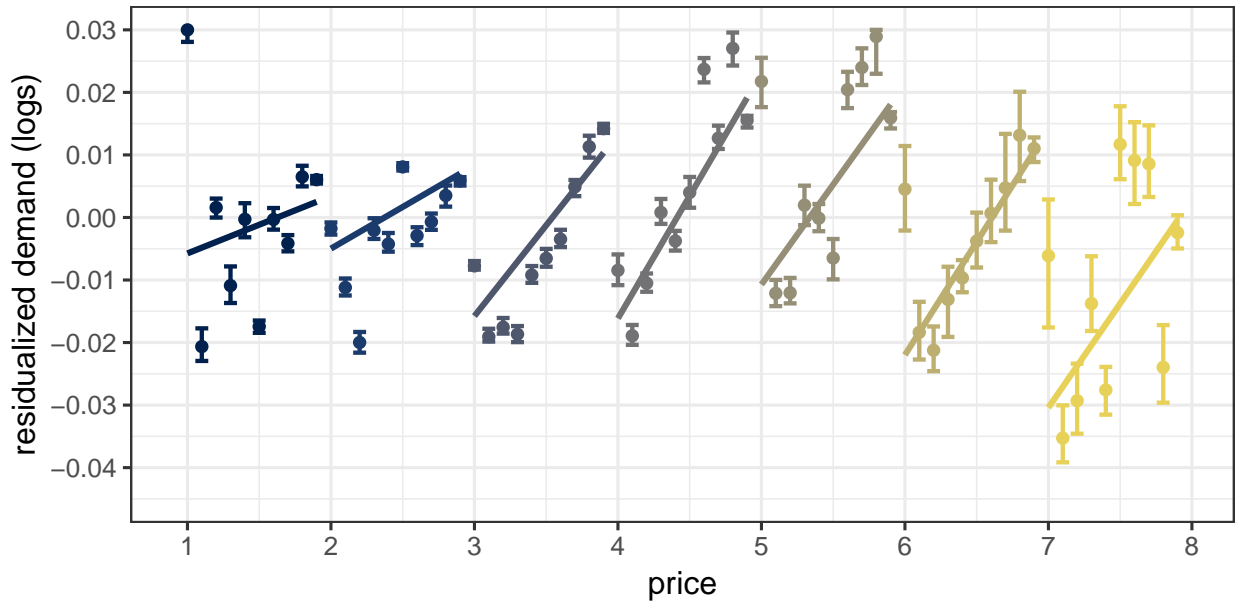
(b) Restricted sample of regular prices, constant semi-elasticity OLS estimation

Robustness: Alternative demand structure. Net log-demand by price, residualizing product-level price *semi*-elasticity, cross-semi-elasticity, seasonality, store fixed effects, and promotion effects. Regressions are conducted at the product level (1710 products, 248 stores, 177 weeks). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Fixed effects are estimated at 10-cent bins, and 95% confidence intervals are result of 200 cluster-bootstraps by store.

The top panel is using data on all to-the-cent prices from a national retailer data (Column (3) in Table 1), and the regression is run with instrumental variables, as in Specification 6 but on prices rather than log-prices. i.e., logQ ~ p, instrumenting for the price with the price of the same item in other stores of the same distribution center in other cities. The bottom panel presents the results of an OLS constructed on regular prices only (column (4) of Table 1). Data is truncated for prices between $1 and $7.99 for clarity. For further details see Section 4.1.

Figure A.6: Sample robustness: Residualized demand for an alternative sample of products
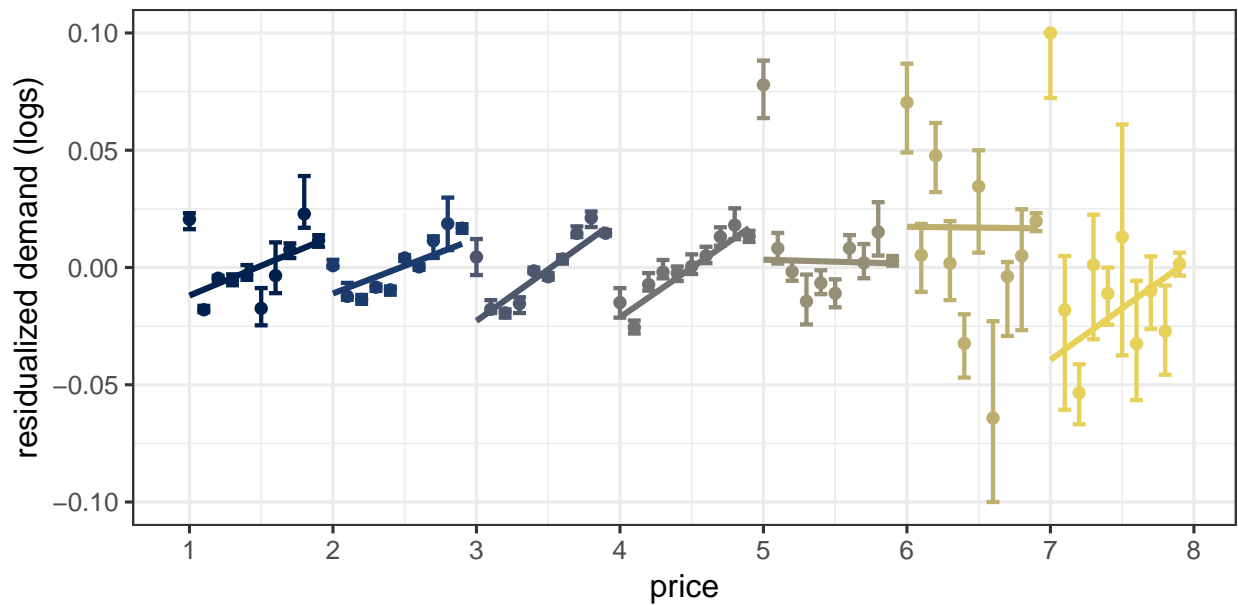
Robustness: alternative sample. Net log-demand by price, residualizing product-level price elasticity, seasonality, store fixed effects, and promotion effects. Regressions are conducted at the product level (2149 products). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Fixed effects are estimated at 10-cent bins, and 95% confidence intervals are result of 100 cluster-bootstraps by store. Top panel shows the entire sample of to-the-cent prices, while the bottom restricts attention to regular prices only.
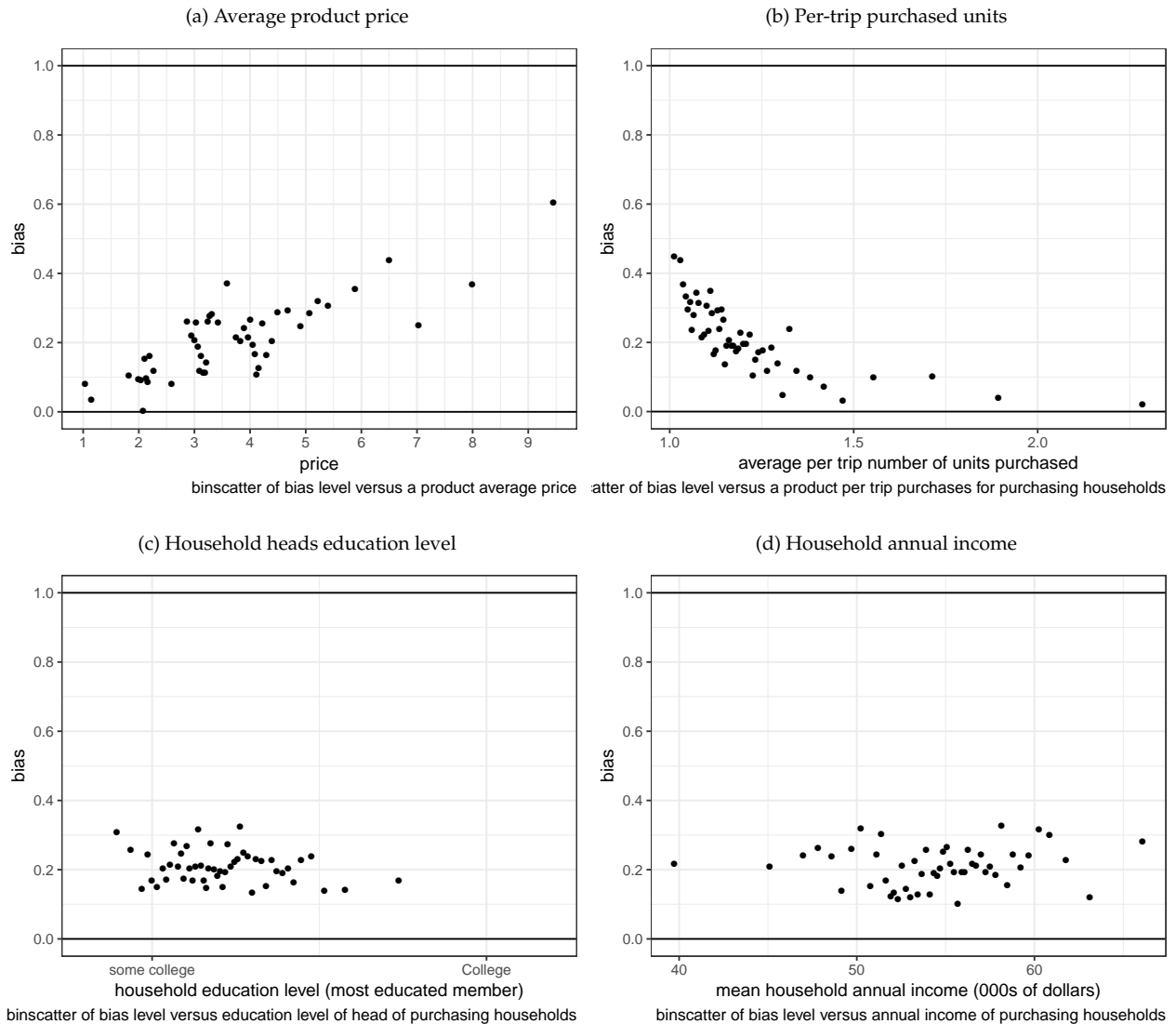
The sample is of 2149 products of high availability but also highly concentrated prices (125 with more than 75% of observations at 99-ending prices, the rest with more than 75% of observations sharing common dollar digit). Sample consisted of 55.5 Million observations spanning the same stores and period of time as for the main sample. Data cover $292 Million in annual revenues via sale of 170 Million units per year. Mean price of these products is $2.20 (1.27), with 21.7% of observations representing on-sale prices.

Figure A.7: CDF of left-digit bias estimates on Retailer data, three estimation methods.
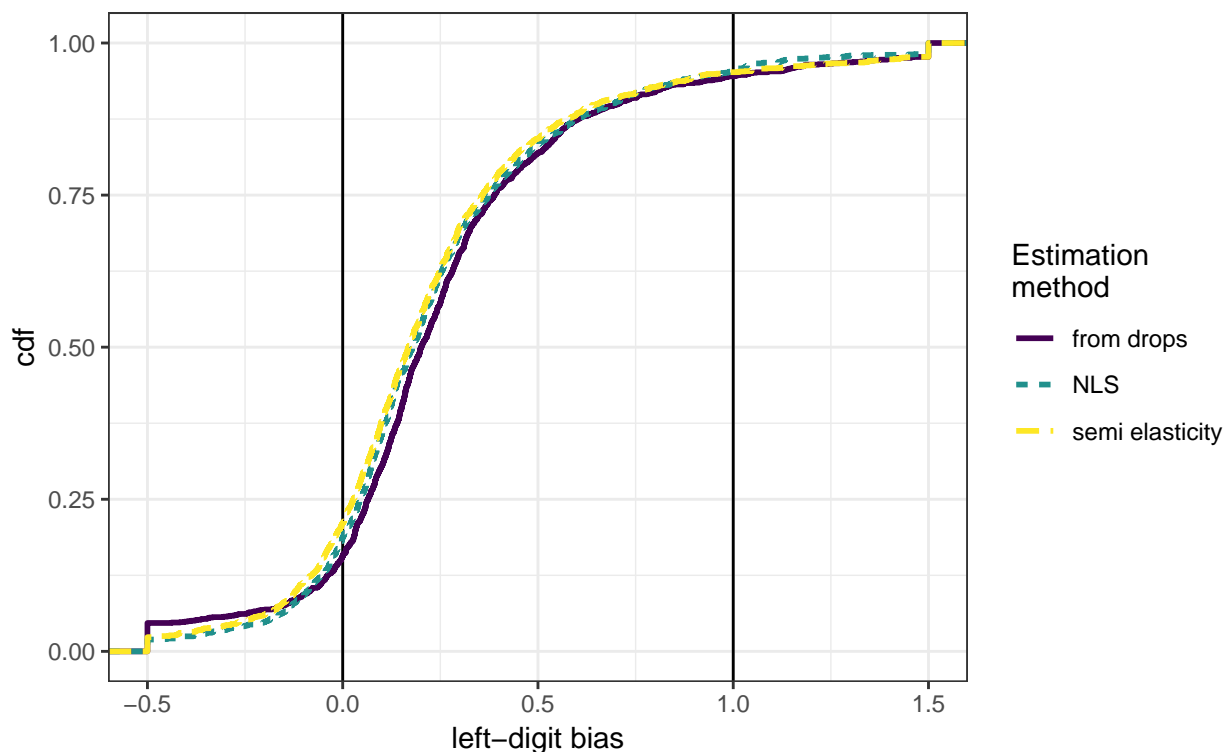
The figure shows the CDF of product-level left-digit bias estimates, corresponding to panel B of Table 2. The dark solid line, "From drops", is the CDF of left-digit bias estimates at the product level using the drops in demand at dollar crossings coupled with elasticity of demand and the dollar digit. Dashed green, "NLS", are estimates from non-linear least squares. Light long dashed, "semi elasticity", are the estimates from a log-linear regression with constant semi-elasticity of demand. For details see Section 4.

Figure A.8: Bias Heterogeneity: Correlates of product and clientele characteristics with left-digt bias

(a) Average product price

binscatter of bias level versus a product average price

(b) Per-trip purchased units

atter of bias level versus a product per trip purchases for purchasing households

(c) Household heads education level

binscatter of bias level versus education level of head of purchasing households

(d) Household annual income

binscatter of bias level versus annual income of purchasing households

The figures show the correlation between product-level left-digit bias estimates and underlying characteristics of the product and its clientele. Product-level estimates are averaged in 50 equally sized bins of the horizontal axis. In each facet, the horizontal axis is the characteristic variable, and the vertical axis is the left-digit bias estimates (from non-linear least squares estimates). Data on clientele characteristics are calculated from Nielsen home panel.

Panels (a)-(d) show the association with the product overall average price, the average number of units a purchasing household purchase of that item in a single trip, the education level of purchasing household heads (maximum level if two heads of household), and the annual income of purchasing households. For further details see Appendix Section E.

Figure A.9: Simulation: Relative price level by bias
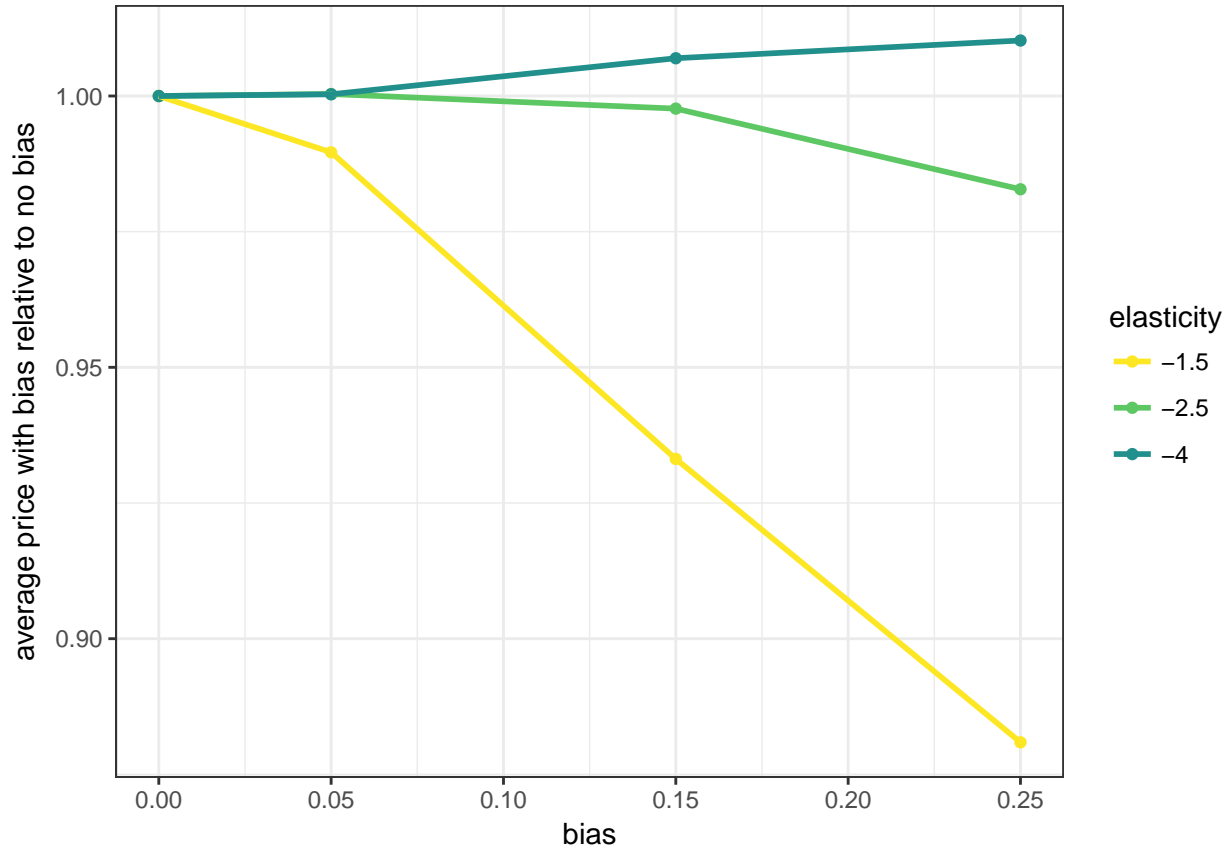
The figure shows the the average price of products set by an optimizing firm by the level of left-digit bias, keeping everything else fixed. The horizontal axis is the degree of left-digit bias, the vertical axis is the relative price to a no-bias case, normalized to 1. Each curve represents a different price elasticity. For low elasticities, the average price goes down with higher left-digit bias because the dominating force is to push prices down to 99-endings. For more elastic demand, the price goes up, because the dominating force is on interior prices, or pushing to a higher 99-ending price.

Figure A.10: Illustration: pricing schedule of a 2-product monopolist
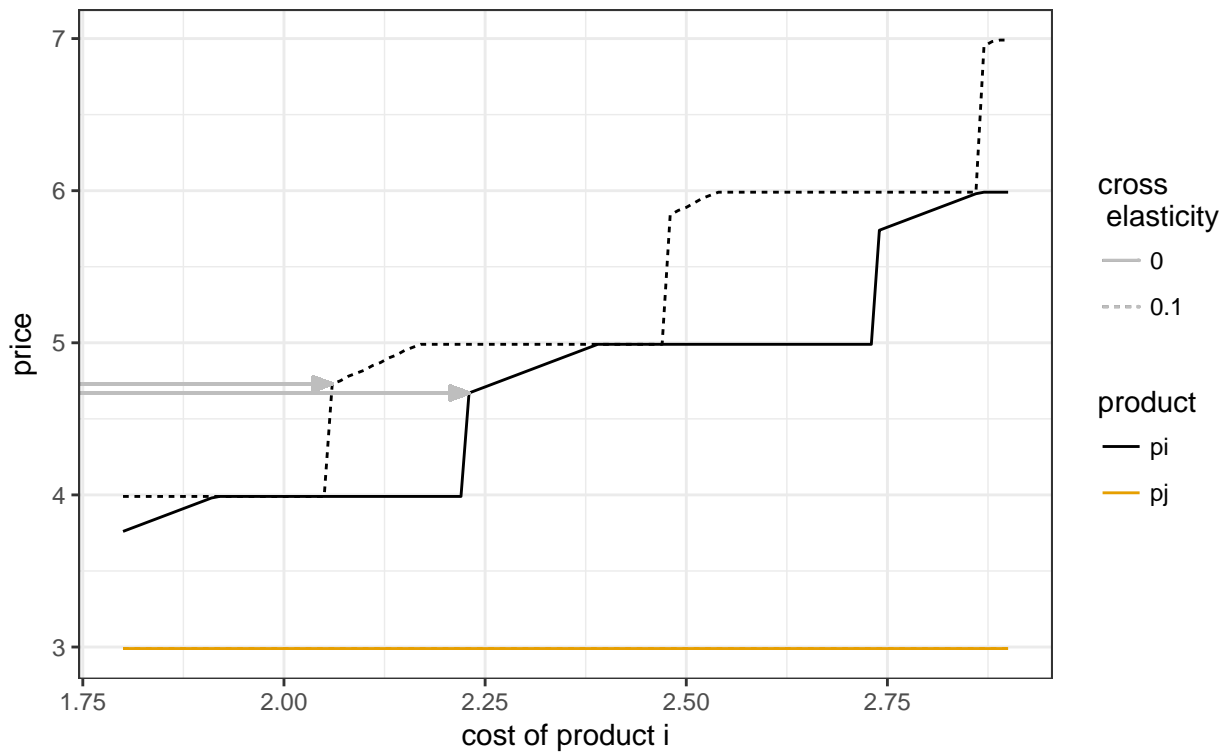
The figure gives an example of a pricing schedule of two products with and without cross elasticities of substitution between the items. The same-price elasticity of each product is -2. Black curves are the prices of product i and orange curves, which lie on top of another, are the prices of product j (the cost of product j is kept constant at 1.5). Solid lines are the prices if there is no substitution between the products, and dashed when there is a cross-elasticity of 0.1. The gray arrows point to the lowest prices of product i starting at 4 in the two different cases, showing that the next-lowest price is higher for substitutes.

Figure A.11: Minimum distance estimation: empricial and predicted moments
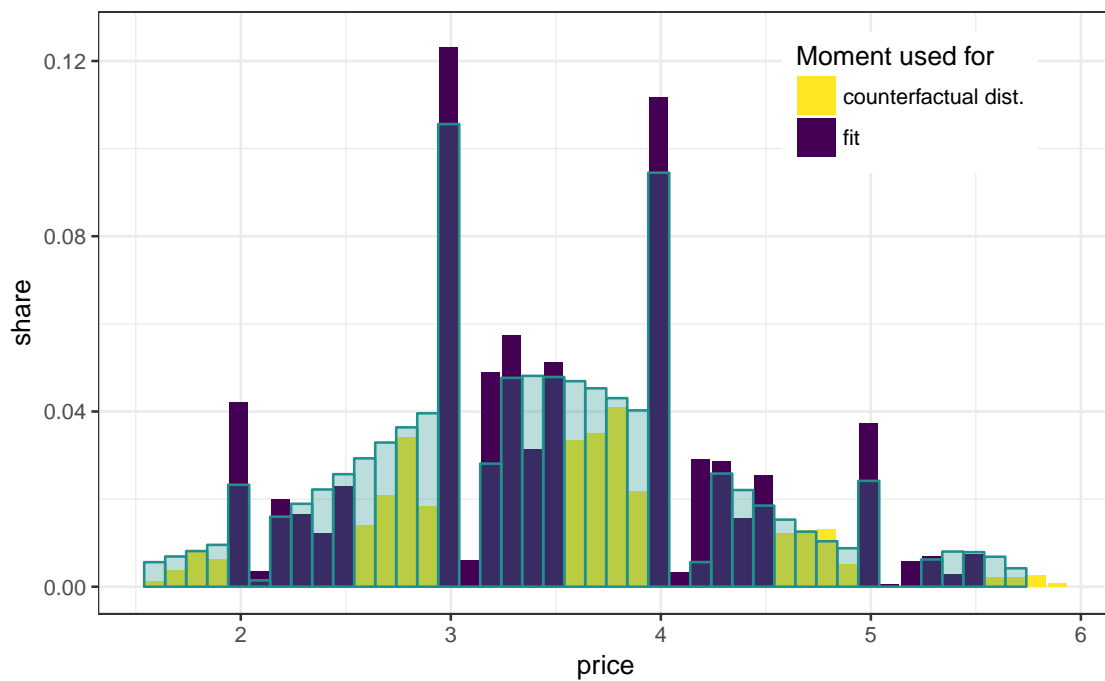
The figure shows the empirical moments and predicted moments from a minimum distance estimation. Each moment is share of prices in a 10-cent bin. The empirical moments are the shares of regular prices for the 1710 products, shown in solid bars. Dark bars are the moments used for fitting, while the light bars are excluded from the fit and used to create the counterfactual price distribution. The predicted moments are shown in semi-transparent light green. For further details see Section 5.

Figure A.12: Estimated firm-perceived elasticity
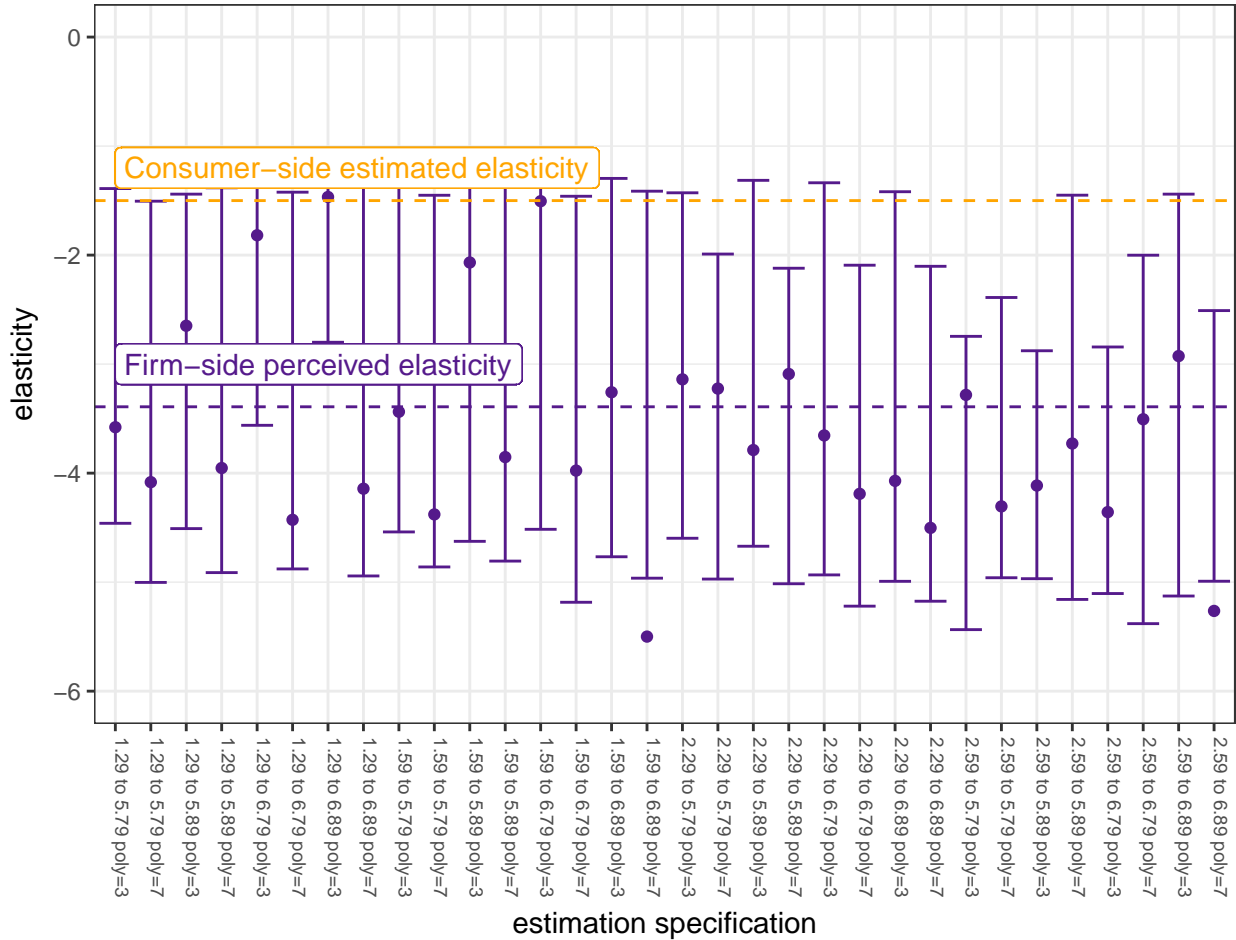
The figure shows the estimated perceived elasticity inferred from the price distribution. On the horizontal axis are different assumptions about the estimation procedure. Namely, the range of prices used for estimation, the excluded prices from the counterfactual price distribution fit, and the shape of that distribution via the polynomial degree. Each point is the estimated perceived elasticity for each specification, and the confidence intervals are the 2.5% and 97.5% estimates from 300 cluster-bootstraps of the moments at the product level (that is why some point estimates are outside of the confidence interval).

For all specifications, I can reject that the elasticity is larger than -1. While point estimates are sensitive to the exact specification, bootstrapping the sample provide more similar confidence intervals.

Table A.1: Summary Statistics and Data Selection, Nielsen Data

| | Nielsen RMS | | |
| | Full | 2+ weeks | 6+ weeks |
	(1)	(2)	(3)
Panel A: Data Description			
Observations (M)	27.8	20.5	11.5
Products	12	12	12
Stores	11000	11000	11000
Weeks	468	468	468
Chains	70	70	70
Annual Revenue ($M)	97.9	57.2	28.4
Annual Units Sold (M)	20.9	11.5	5.67
First Date	Jan-06	Jan-06	Jan-06
Last Date	Dec-14	Dec-14	Dec-14
Panel B: Pricing Descriptives			
Mean Price	5.51	5.55	5.32
Price SD	2.81	2.78	2.57
Share on-sale			
Share 99-ending	0.30	0.35	0.34
Share 9-ending	0.70	0.81	0.86
Share 9-ending (non-sale)			
Share 0-ending	0.10	0.09	0.04
Share 0-ending (non-sale)			

The table presents summary statistics for the Nielsen data. Column (1) is the full sample as taken from DellaVigna and Gentzkow (2017). Column (2) is the main sample used, excluding what are likely non-real prices (i.e., keeping only prices that are part of an at least 2-week spell). Column (3) provides further information on the characteristics of longer price spells (at least 6 weeks).

Table A.2: Correlations between product-level estimated bias and product characteristics

Dep: Left-digit bias

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(average price)	0.195*** (0.018)				0.141*** (0.020)	0.157*** (0.030)	0.210*** (0.036)
log(per-trip units)		−0.489*** (0.045)			−0.332*** (0.050)	−0.396*** (0.076)	−0.220** (0.089)
log(income)			0.077 (0.087)		0.129 (0.104)	0.193* (0.113)	0.095 (0.114)
education (z-score)				−0.011 (0.008)	−0.022** (0.010)	−0.019* (0.010)	−0.016 (0.010)
Fixed Effects						Category	SubCategory
Observations	1,504	1,504	1,504	1,504	1,504	1,504	1,504
R^2	0.094	0.076	0.001	0.002	0.123	0.214	0.306

The table shows results from regressing product-level left-digit bias estimates on a product average price, and purchasing households per-trip units bought, income, and education (highest among household heads). Columns (1)-(4) are the univariate correlates, while (5)-(7) include all variables. Column (6) includes 47 category fixed effects (e.g. "coffee/tea/hot cocoa/cream") and column (7) includes 121 sub-category fixed effects (e.g. "coffee"). So in column (7) the variation is between products of the same sub-category.

Table A.3: Profits lost when the firm underestimates left-digit bias, alternative assumptions

	Assumed scenario			profits lost pricing
	Left-digit bias θ	Price elasticity ϵ	Perceived bias $\hat{\theta}$	by perceived bias (%) ($\hat{\theta}$ instead of θ)
Panel I: If there is also left-digit bias to last digits $\theta_2 = 0.25$				
Point estimates	0.16	-1.5	0.004	-2.8
Extreme estimates	0.10	-4.0	0.03	-1.0
Heterogeneous bias	$N(0.16, 0.05)$	-1.5	0.004	-3.2
Heterogeneous perceived bias	0.16	-1.5	$N(0.004, 0.01)$	-2.3
Heterogeneous elasticity	0.16	$N(-1.5, 0.2)$	0.004	-2.8
Panel II: If consumers round up when ignoring the last digits				
Point estimates	0.16	-1.5	0.004	-1.6
Extreme estimates	0.10	-4.0	0.03	-1.0
Heterogeneous bias	$N(0.16, 0.05)$	-1.5	0.004	-1.4
Heterogeneous perceived bias	0.16	-1.5	$N(0.004, 0.01)$	-1.2
Heterogeneous elasticity	0.16	$N(-1.5, 0.2)$	0.004	-1.1

The table shows calculated percentage of gross profits lost, relative to optimal pricing. Optimal means that the firm perceives the bias to be what it actually is, $\hat{\theta} = \theta$, while the other cases take a lower perceived bias into account. Panel I shows the profits lost if the perceived price is $\hat{p} = (1 - \theta - \theta_2) p + \theta p_1 + \theta_2 (p_1 + p_{0.1})$, that is, with extra ignorance of the cent digit only, assuming the firm prices as if $\hat{\theta}_2 = \theta_2 = 0.25$. Panel II shows the profits lost when perceived price is $\hat{p} = (1 - \theta) p + \theta (p_1 + 0.99)$, meaning that consumers assume the price ends with 99 with weight θ. Each row makes a different assumption about the "true" parameters and the perceived ones (but always assuming that the firm knows the true price elasticity).

The first row in each panel shows the lost profits under point estimates taken from the main analysis. The second row, extreme estimates, makes a firm-forthcoming assumption with lower bias, high price elasticity, and the upper bound of perceived bias. The next three rows explore the role of heterogeneity and how consequential it might be: I examine heterogeneity in left-digit bias, in perceived left-digit bias (truncating from below at 0.001), and in same price elasticity. Finally, the last row takes cross-elasticities into account (I use the implied perceived bias of 0.004, even though the implied perceived bias with cross-elasticities, given pricing, is lower. This way I am consistent and provide a lower bound on losses).

At each cell I simulate 3000 costs from a cost distribution that mimics the actual distribution inferred from a constant markup prices. For the heterogeneous cases I draw 50 values for each parameter. For the last row, I sample 80 prices for each of 2 products (from the same cost distribution as above).