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

Avner S. Shlain*
University of California, Berkeley

May 9, 2018

Abstract

Why do so many prices end with 99? Arguably because of left-digit bias, the tendency of consumers to perceive a $4.99 as much lower than $5.00. I solve and estimate a new model of biased consumers and responding firms that makes two key predictions: (1) prices should bunch at 99-ending prices; (2) these prices are bunching from above, with ranges of missing prices with low price-endings. This paper is the first to provide robust reduced form evidence and structurally estimate left-digit bias in a retail setting, by using large scanner data from two sources. I find on the consumer side a bias on the order of 25%, interpreted as consumers ignoring 25 cents of a 99-ending price. Qualitatively, firms respond to the bias, with high shares of 99s and missing low-ending prices that increase with the dollar digits. However, quantitatively, firms act as if the bias is orders of magnitudes smaller, therefore foregoing 1-3 percentage points of profits.

*Shlain: UC Berkeley, shlain@econ.berkeley.edu.

I am thankful for Stefano DellaVigna for endless encouragement and support. Ned Augenblick, Zarek Brot-Goldberg, Ben Handel, Ganesh Iyer, Ulrike Malmendier and Dmitry Taubinsky provided many thoughtful comments. I have also immensely benefited from discussions with Eric Avis, Natalie Cox, Daniel Haanwinkel, Yoram Halev, Johannes Hermle, Peter Jones, Shachar Kariv, Kei Kawai, Jennifer Kwok, Julien Lafontaine, Nick Li, John Loeser, Filip Matejka, Waldo Ojeda, Devin Pope, Matthew Rabin, Gautam Rao, Alex Rees-Jones, Jon Schellenberg, Yotam Shem-Tov, David Sraer, Alex Steiny, Justin Sydnor, Jonas Tungodden, Miguel Villas-Boas, and probably many others who I hope will forgive my imperfect memory or forgot themselves. I am thankful for Liran Einav, Nir Jaimovich, Saul Lach, Rony Rojkin, Maya Segal, Sofia Villas-Boas, and Merav Yiftach in assisting me with data collection. Most of all, I am thankful for my spouse, Osnat Strulov-Shlain, who sent me a link to an online article saying “might be interesting for research” a few years ago. Certain sections were calculated (or derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

This work has been supported (in part) by award # 98-18-04 from the Russell Sage Foundation. Any opinions expressed are those of the author alone and should not be construed as representing the opinions of the foundation.
1 Introduction

Why do so many prices end with 9, and in particular with 99? This is likely due to left-digit bias, leading
to drops in demand when left-most digits of price change. The bias has strong implications on demand,
and hence on pricing and profits. The existence and estimated magnitude of the bias leads to increased
profits (for the optimizing firm) of a few percentage points compared to selling to non-biased consumers.
However, firms price as if the bias is existing but much lower, and lose most of these gains.

Previous papers have explored the 9-ending pricing phenomenon, suggesting different theories and pro-
viding various kinds of empirical support. However, no paper had empirically explored consumer bias
and firm response jointly. Therefore, no paper had quantified the bias, nor discussed the quantitative im-
plications of the bias on firm behavior and profits. While most papers provide some support of the bias
existence, it is unclear how common it is, unknown what is its magnitude, and therefore unclear if it matters
for consumers or firms.

This paper is the first to take a model of both sides of the market, consumers and firms, and estimate it using
big data. Doing so, I document robust evidence for the bias existence for thousands of grocery products,
structurally estimate it, and discuss its underlying mechanisms, implications, and firms response to it. I
conclude by conducting the first counterfactual exercises of the effects of the bias and firm pricing behavior
on profits.

There is strong sorting for 9- and 99-ending prices, including in the two scanner datasets I use. I am using
a sample of 1710 popular products in 250 stores of a single US retailer over 3.5 years, and 12 products in
AC Nielsen RMS data across 60 chains, 11,000 stores, over 9 years. I argue that the 99 pricing behavior,
which holds across firms, countries, products, and time, is an indirect revealed preference evidence for a
bias exploited by firms. I model this bias as if consumers drop off $\theta$ share of the lower digits of a price,
leading a price $p$ to be perceived as $\hat{p} = (1 - \theta) p + \theta \lfloor p \rfloor$ (a mix with weight $\theta$ of the dollar digits, i.e. the
price with no cents, and $1 - \theta$ of the exact price). The bias causes aggregate demand to drop when the first
digit changes, and also dampens the price sensitivity to within-dollar price changes.

Next, I turn to explore whether the bias is evident on the demand side. Even though firms are strategic
and price endings are endogenous, if the model is correct, we should observe drops in demand crossing the
dollar digits. As long as there is some variation in price endings, we can estimate the demand structure. In
practice, under the model, residuals by price from a logQ-logP regression display a seesaw pattern - drop-
ning across dollar thresholds, and increasing within each dollar-digit. Several challenges exist when trying
to get clean evidence of residuals by price in practice. The main challenges are (1) that many observations
in scanner data are average prices that no consumer actually paid in practice, and are thus uninformative
about the response to an exact price; (2) that some price endings are also tied to specific promotion strate-
gies, acting as demand shifters; (3) that price endings are chosen endogenously, so a simple comparison
between $4.99$ and $5$ is impossible to do; and (4) that as with every attempted demand estimation we
might worry about simultaneity and price endogeneity. I find a novel transparent way to deal with measu-
rement error in scanner data by restricting attention to what are likely to be observations of “real” prices
(prices consumers actually paid), and to proxy for sale behavior by using the length of price spells and the
last digit of a price. I argue that the price-ending endogeneity is not biasing the demand estimation, just
make it noisier. I use multiple fixed effects to control for most endogeneity worries, and supplement it with
a form of Hausman Instruments with some added benefits from the richness of the data. Following these
data cleaning procedure and estimation technique, the drops in demand are visible in the data. That is, it is
visually observed that demand is flatter across cents within a dollar-digit and is dropping when the dollar
digit change. For example, a change in price from, say, $4.99$ to $5.00$, is tied to a drop in demand of about
4-5 percentage points in number of units sold. These drops, together with the price level and price elasticity, can then be translated to approximate the bias parameter. Namely, consider a drop of 4 percentage points in demand crossing from $4.99 to $5.00, and an estimated price elasticity of -2. Then, a change of 4% in demand can also be coming from a 2% change in price. 2% of $5 are 10 cents, implying a bias of 0.1, since for these parameters demand is the same at $4.90 (≈ (1 − 0.1) * 4.99 + 0.1 * 4) as it is for $5.00. I then estimate for each product the drops in every dollar digit, and its own price elasticity (using instrumental variables), to calculate the bias. To estimate the bias more precisely, I estimate the model directly with non-linear least squares at the product level. 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 very similar estimated parameters, with rank correlation of 0.94 between the NLS and log-linear estimates, and 0.7 between the drops-based estimates and log-linear estimates. The average bias level is 0.22-0.25 depending on the estimation technique. Meaning, consumers ignore about 25 cents of a 99-ending price.

The bias estimates by product can be considered outcome variables, and allow to ask what product and clientele characteristics are correlated with higher bias. If the bias is driven by relative or proportional 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, the relation to price might be misleading if higher ticket items are 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. Along these lines, more educated people might be expected to exhibit less bias (e.g., as in Bronnenberg et al. (2015)). Finally, the effect of income is interesting even if ex-ante unclear (higher income might mean lower sensitivity to price overall, while low income is often correlated with higher behavioral biases). Calculating for each product measures of these four variables, and regressing the bias on all of them, with added product controls, show correlations with the expected signs. Overwhelmingly the strongest and most robust relation are to price and units purchased. Namely, a 10% higher priced item is correlated with a 0.021 higher bias (meaning, ignoring additional 2.1 cents out of a 99 ending price), and an item purchased with 10% more units per trip is correlated with 0.022 lower bias. Income is positively correlated with the bias, and education is negatively correlated, but not in a robust way.

The first part shows that the bias exists and suggests some mechanisms of where it is coming from. Now we can turn to the pricing firms and ask - (1) how should firms price? (2) do they follow that pricing prescription? If not, (3) what are the implied perceived bias and elasticity levels they consider to exist? and (4) what are the possible implications on firm’s profits?

The model predicts the price given bias, elasticity, and cost. Discontinuities in demand lead to a pricing schedule by monopolistic firms (or competing firms, in Appendix Section B) that has two distinct features. 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 from above. That is, there will be a well-defined region of missing prices above the 99-ending price. While the prevalence of 99-ending prices has captured the attention of many researchers, the latter prediction had not been explored before, and separates this model from the few other models suggested in the literature. Indeed, not only do retail prices end with 99 (24%-34%) but they rarely end with a price ending lower than 19. When they do, it is likely to be an artifact of other 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, for $3.99 with a bias of 0.05 and elasticity of -2, the next-lowest price is $4.65. The next-lowest price above $4.99, i.e., starting with $5, is $5.69. This is corroborated in firm pricing
behavior, as the lowest percentiles of price-endings are increasing as the dollar digit increases. However, quantitatively, there is a strong discrepancy between the estimates and firm pricing. The estimated levels of left-digit bias (0.25 or even 0.15) and price elasticity imply that all prices should end with 99. More elastic demand, or lower bias, decrease the next-lowest prices and hence in principle, can reconcile the patterns in the data as if the firm perceives the bias as lower than it is, or demand more elastic. However, simple calibration exercises show that even for unrealistically high elasticities (such as -15, an order of magnitude higher than estimated), 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. Indeed, pricing is consistent with a much weaker perceived bias level (more than an order of magnitude weaker).

Finally, armed with demand side estimation, one can ask what are firm profits under different counterfactual exercises. Under the model, the mere existence of the bias, even if firms ignore it, lead to higher profits by a few percentage points, simply because demand is higher almost everywhere\(^1\). If firms price optimally in response to the bias, profits are higher by another few percentage points. For example, with the estimated elasticity and bias, profits are higher by about 4.5%. However, if firms price as if the bias is much weaker, they lose most of the possible gains. These losses are driven by pricing at dominated prices, thus not exploiting the excess distorted demand.

Of course, the prevalence of 9 ending prices is not new, and has been extensively explored (see Ginzberg (1936) for an early example). The literature has focused on documenting and testing two main classes of explanations – “level effects”, by which the last digits of a price are ignored, and “image effects”, where the price ending serves as a signal regarding the product’s quality or value (see Stiving and Winer (1997) for a review). Empirical and experimental findings are mixed, but generally support the existence of an effect (e.g. Anderson and Simester (2003); Bizer and Schindler (2005); Carver and Padgett (2012); Hackl et al. (2014); Snir et al. (2017); Stiving and Winer (1997); Thomas and Morwitz (2005)). Practically all papers are documenting direct or indirect evidence for existence of the bias, and only few of them do so for retail pricing where the bias is most relevant (Stiving and Winer (1997); Macé (2012)). As argued above, this paper is the first to provide large scale reduced form evidence. On the theory side, some of the few models that do exist (e.g. Basu (1997, 2006); Gabaix (2014)), give predictions that are challenging to reconcile with data, while a smooth model of left-digit bias (DellaVigna (2009), similar to the one used here on the demand side) was estimated in various economic settings, but not on prices (Lacetera et al. (2012)). This paper is the first to estimate a structural model of the bias, and the first to explore actual and predicted firm responses.

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

2 Model

existing theories Most of the literature on 99 pricing is not formally modeled, but it is assumed the firms do so in response to some non-standard response of consumers to pricing. Main explanations are in the

\(^1\)This prediction relies on a strong assumption, that perceived prices are attenuated downward due to the left-digit bias. If instead they are attenuated up, then of course profits will be lower with the bias, since demand will be lower everywhere. Some experimental papers support the former interpretation, but the evidence is not solid.
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. That is, Basu’s model predicts that price distribution is either collapsed to 99-ending prices, or that 99 plays no role. This is an interesting model with intuitive appeal, but it leads to equilibria that are not be supported in the data. Similar take, but more flexible, is taken in (the online appendix of) Gabaix (2014) where attention is also endogenous, and a monopoly chooses to price at the default price for a variety of marginal costs. A key distinction between Gabaix’s model and the model I will be introducing shortly, is that Gabaix predicts missing prices below the “bunched” price, but not above it. This stands in contrast with the data. Chen, Iyer and Pazgal (2010) model a game between consumers and competing firms, where consumers are “price categorizing”, i.e., can not exactly recall a price but can use a finite partition and recall to which element the price belongs. This model explains a discrete price distribution, but it includes further sophistication at the consumer level in choosing categories. The category choice comes to maximize utility, and is unrelated to the price ending. In other words, these models are not suitable to generate a specific price ending, such as 99.

A simple model of left-digit bias was suggested by DellaVigna (2009). The model was later estimated by Lacetera, Pope and Sydnor (2012) by examining the effect of odometer readings on prices in the used cars market. This model of consumer behavior, which I will soon complement with a responding pricing firm, posits that each consumer reads a number left to right, and has a probability $\theta$ of ignoring the rightmost digits. In Lacetera et al. (2012), that number is the odometer of the car – an inherently different setting than retail, since odometers are harder to manipulate (and indeed the authors do not find strong selection across thresholds).

I will now introduce the model. It consists of left-digit-biased consumers in the spirit of DellaVigna (2009) and monopolistic firms. The bias causes aggregate demand to be discontinuous when the left-most digit change, in turn leading firms to bunch at prices just below.

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 observe the price as possibly different than it actually is. For example, in Chetty et al. (2009) we can think of $\hat{p}$ as the price with inattention to sales tax, and in Gedenk and Sattler (1999) as the price rounded to whole dollars or floored to the nearest dime. 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. One interpretation is that it describes that each consumer is averaging - perceiving the price in full $1 - \theta$ fraction of the time and only the first digits $\theta$ fraction of the time. A second alternative is creating a type mixture. Namely, there is a share $\theta$ of consumers who only observe the dollars digits and a share $1 - \theta$ who observe the price in full. Though different in interpretation, these two modes lead to similar qualitative predictions. Since the intrinsic inattention is slightly nicer to deal with arithmetically, this will be the model of choice. That is, from now on assume that

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

(1)
A crucial modeling choice made here is that if lower digits are not observed, they are completely dropped. This is important for welfare calculations (do consumers over- or under-consume?), but can not be identified in the data. Since this is the simplest formulation, and the one used in other papers, I will use it from now on. However, I will add different robustness tests when this assumption is consequential.

If the agent has utility of the form \( U(y, q) = y + \frac{\theta + \frac{1}{2}}{1 - \theta} \) 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

\[
D(p; \theta) = A p^{\theta} = A ((1 - \theta) p + \theta \lfloor p \rfloor)^{\theta}
\]

To get some intuition, the impact of this bias parameter (in other papers titled “inattention”) can be understood by observing three distinct consequences for price setting and demand: First, 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 \( \theta \) is higher; Second, 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); Last, the penalty of a 0.01 price increase is higher if higher level digits also change (i.e. 2.99 to 3.00).

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 profits are

\[
\Pi(p; \theta) = D(p; \theta)(p - c)
\]

e.g. 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.

### 2.1 solution

Assume internally inattentive demand as in\(^2\) Profits are \( \Pi = D(p; \theta, c)(p - c) \). For ease of exposition, I decompose a price \( p \) into its decimal basis components, s.t. \( p = p_{10} + p_{0.1} + p_{0.01} \). For example, if \( p = 3.49 \) then I write it as \( (3, 0.40, 0.09) \).

With left-digit bias, we get a discontinuity in demand when changing \( p_{1} \). To simplify the solution, assume that prices can be chosen from \( \cup q_{1} q \subseteq R_{+}\) where \( q_{1} \in N \) and \( q = (q_{1}, .9, .09) \), i.e. a subset of the positive real numbers with all numbers between a natural number and a .99 ending number. Proofs to the following propositions and corollaries appear in Appendix A.

This leads to the following pricing schedule for small \( \theta \)\(^3\)

**Proposition 1. Optimal pricing formula, for small \( \theta \).**

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

\[
p(c; \theta, \epsilon) = \begin{cases} 
q & \text{if } c \in \left[ L_{q}, L_{q+1} \right] \\
\left( c - \frac{\theta q_{1} + 1}{\epsilon} \right) \frac{1}{1 + \epsilon} & \text{if } c \in \left( L_{q}, L_{q+1} \right)
\end{cases}
\]

\(^2\)Conversely, if there is type mixture as described above, then following the same formulation will lead to demand of the form \( D(p; \theta) = A ((1 - \theta) p^{\theta} + \theta \lfloor p \rfloor^{\theta}) \).

\(^3\)For example, an upper bound on \( \theta \) is that \( \theta < 1 + \frac{1}{\epsilon} \), otherwise the range of interior solution is empty, as can be seen from the definition in Proposition.
Where $c_q$ and $c_{q+1}$ 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 $q$.

This pricing behavior is different than the no-bias case in two ways. First, prices bunch at 99, and second, interior solutions are 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 intuition for the first region is that the drops in demand when changing the left-most digit of the price lead to lower profits in the interior solutions than in sticking with 99. Thus, this model generates price stickiness that is not driven by frictions on the supplier side.

The lower expression in Equation 3 are the interior solutions. It describes the optimal price as a modified Lerner equation with an extra term $\left(-\frac{\theta}{1+\theta} \cdot \frac{q_1+1}{q+1}\right)$. In the interior solutions region, the price is slightly higher than it would have been absent the bias (recall that $\epsilon < -1$). The gap is increasing in the bias ($\theta$), and the nominal price ($q_1$), and decreasing in the absolute elasticity ($\epsilon$). The pricing schedule is illustrated in Figure 1 where the diagonal gray line is the price without bias, the purple line describes the interior solutions, and the green horizontal lines are the 99 ending prices.

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.

To further explain Equation 3 we can discuss the three threshold costs. $c_q$ and $c_{q+1}$ 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 $q + 1$ ((or graphically, where the isoprofit is tangent to the demand curve). The image of costs in $[c_q, c_{q+1}]$ is hence prices in $[q, q + 1]$. The third threshold, $\bar{c}_q$, is the most interesting one. This is the cost for which prices will switch between $q$ to a point in $[q_1 + 1, q + 1]$. At $\bar{c}_q$ two conditions are met: (1) the profit is maximized (with an internal solution) somewhere on the segment above $q$, i.e. on some price $P_q \in [q_1 + 1, q + 1]$, and (2) profits are equal at that price, $P_q$, and at $q$. For ease of notation, I omit $q$ from $P_q$ in the following:

$$\bar{c}_q = \left\{ \begin{array}{l}
P + \frac{(1-\theta)P + \theta P_1}{\frac{\epsilon(1-\theta)}{D_P - D_{\bar{c}_q}}} \\
\end{array} \right. \quad (4)$$

Giving one implicit function for the next-lowest price $P$ as a function of the parameters $\theta$ and $\epsilon$:

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

This equation is not analytically solvable, but can be easily solved numerically. Given $P$, we get $\bar{c}_q$ from Equation 4.

**Comparative statics** We are interested in the effect of the elasticity and inattention on the next-lowest-price $P$, and can show the following results:
Proposition 2. Comparative statics of the NLP: 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. $\frac{\partial p}{\partial \theta} > 0$.

More elastic demand leads to stronger decline in demand as the price changes. The proof is not insightful, and for good reasons. Higher elasticity leads to competing effects on the next-lowest price. Higher elasticity means larger drops in demand (since a price lower by $\theta$ cents is tied to lower demand), but also that margins are lower everywhere, including at the next-lowest price. Overall, it leads to lower next-lowest prices. The effect of higher bias is simpler, since demand is rotated out, and the demand drops are larger, both pushing the next-lowest prices up.

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. 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. Note, that apart from having a connected support in $[p^{-1}(q), p^{-1}(P)]$ (i.e., 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 3. The lower digits component of the Next-Lowest Price (i.e., $P - P_1$) are increasing with the first digits, for $\epsilon < -1$.

The structure of $F$ is crucial for the excess-mass predictions, and hence less robust. But, we can show the following regarding the excess mass at 99-ending prices:

Proposition 4. Comparative statics of excess mass: If the cost distribution is locally uniform, then

Excess mass at $q$ is higher when bias is stronger, i.e. $\frac{\partial \Delta q}{\partial \theta} > 0$

Excess mass at $q$ is changing ambiguously for higher elasticity.

The intuition behind the effect of bias is straightforward - a stronger bias is like an outward rotation of the demand curve, pushing more prices to $q$ both from below and from above. However, the effect of the price elasticity is trickier. Consider a uniform cost distribution. In a bias free world, if $c \sim U[a, b]$ then $p \sim U \left[ a \left( \frac{|\epsilon|}{|\epsilon|-1} \right), b \left( \frac{|\epsilon|}{|\epsilon|-1} \right) \right]$. So, a higher elasticity means a more condensed price distribution. Add to that the bias, causing bunching, and the first order effect will be larger mass at any discrete price. However, the elasticity also reduces the NLP, thus less prices from above $q$ are now shifted to $q$. Which effect dominates relies on the elasticity level. If the cost distribution is locally uniform (meaning a linear slope of the CDF), and if demand is more elastic than 2 (which is usually the finding for retail products), excess mass will increase.

Joining Propositions 2 and 4 means that one can identify the parameters of elasticity and bias from the two moments of excess mass at $q$ and next-lowest price above $q$.

3 Literature review

9 is overwhelmingly the most common price ending, especially in the US. Previous papers document 99-ending prices in shares of 15% to more than 90%. While there are different explanations and names for this

[1]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 to the heterogeneous-bias Taubinsky and Rees-Jones (2016)).

8
phenomenon, they all share in common the trait that the reduction in demand by increasing a price from, e.g. 2.98 to 2.99, is much smaller than the reduction in increasing the price from 2.99 to 3.00. I am using the term “left-digit bias” to describe this trait.

In this section I review the evidence supporting the existence of a bias in several forms: direct evidence from previous field experiments showing demand effects of the 9 price ending; lab evidence for the underlying mechanisms of drop-off or left-to-right processing; discrete price distribution with bunching at 9-ending prices as a revealed preference of pricing firms favoring these price endings over others; and finally, that firms revert to the 9-ending pricing after denomination reforms - in the Euro zone and Israel. I will conclude with possibly alternative explanations for the pricing evidence and argue that in fact they do not offer an alternative.

3.1 Direct evidence

Direct evidence for the bias is actually hard to come by in observational data. The reason is simple. If some prices are better than others, pricing firms will select their prices accordingly. To the extent that sorting is strong but not perfect (for example 45% of prices end with 99, but there are also 2% of 00 in one of my data sets), one must ask what is special about the “sub-optimal” observations. In other settings, this is not as much of an issue as it is in retail pricing. For example, the paper by Lacetera et al. (2012) documents left-digit bias regarding odometers of used cars. There, manipulation of the mileage seems to be very limited. In contrast, Repetto and Solis (2017) and Chava and Yao (2017) document significant sorting in posted house prices in Sweden and the US (respectively). Repetto and Solis (2017) indeed spend most of their analysis on showing that there is no clear sorting on observables, and control for as much as they can, but one can still think of the (very large) effects of the listing price on final sale price as a lower bound. As I will show below, and as was found previously, in retail pricing there is very strong sorting, making the simple direct comparison – such as demand at $2.98 versus $2.99 versus $3.00 – almost impossible to conduct. There is still a way, that I will introduce in the next section

Of course, the ideal setting will be a randomized controlled trial in stores, where the price of a good is set to end at different price points. Randomization should be enough to allow estimation of demand slopes below and above a round number, such that it is feasible to separately identify a discrete jump when going from price endings of 99 to 00. To the best of my knowledge, only a few attempts are documented in academic writing, mostly in direct-mail catalogs (Schindler and Kibarian (1996); ?) with one exception of in-store experiment (Bray and Harris (2006)). There were also precedents from earlier days, but with statistical ambiguity (Ginzberg (1936); Dalrymple and Haines Jr (1970)). These papers provide much insight – most of these papers find positive effect on demand of 9-endings prices, while Bray and Harris (2006) find a negative effect. However, there are some limitations to each, and all papers have limited demand volume. A clean experiment, is by Anderson and Simester (2003). They run a set of 3 field experiments where they randomize prices for women fashion items sold through a mailed catalog. Some key differences from the supermarkets setting, is that the direct-mail catalogs are for high-priced items (mean price of $58) and all prices are in whole dollars (so think of comparing $58 to $59 to $60). The 3rd experiment contains the needed variation to pick up this discrete jump, but their specification is not addressing this notion directly (rather, they test if there is higher demand for $9 ending prices per se), and the within product price variation is at most 3 prices, not necessarily above and below a “threshold”. Bray and Harris (2006) conducted a field experiment on 10 products in 12 stores of a UK retailer (6 control stores ending with .X9, 6 treatment stores ending with .X9, 5 further “image effects” explanations will argue that there is an increase in demand by going from 2.98 to 2.99, since 9 carries a special signal about value.
For 9 of these products, sales were higher for the round price. However, serious concerns over the statistical test and validation of the treatment/control assignment are not addressed in the paper.

Another example is a working paper by Ashton (2014). Ashton examines the data from Chetty, Looney, and Kroft (2009). Chetty, Looney, and Kroft have experimentally varied inclusion of the sales tax in price tags in several stores, and their paper focuses on the salience of the tax rate. They find that, on average, consumers are under-attentive to the sales tax. Ashton is showing in reduced form, that most of the demand reduction due to the posted tax is coming from instances in which the tax inclusive price is also crossing the dollar threshold. This is, again, supporting the notion of a large drop in demand when changing the first digit, but not otherwise (importantly, for the same percentage increase in price).

3.2 Lab evidence

Direct mechanisms were examined in experimental papers. A variety of papers focused on the recall accuracy and mean error by price ending (e.g., Schindler and Wiman (1989); Schindler and Kibarian (1996); Schindler and Kirby (1997); Schindler and Chandrashekaran (2004); Snir et al. (2017); ?; see summary table in Carver and Padgett (2012)). Taken together, these papers find that people are worse in recalling prices ending with 99 than prices ending with 00. Finer level findings (as the hazard depending on the exact price ending) are harder to measure. One exception, also related to a reform analyzed in a companion paper (Shlain (2018)), is by Snir et al. (2017). They find that 99-ending prices are more likely to be recalled with error, and also find a positive, yet not statistically significant, higher recall error for any non-0 price ending (Table 2 column 3).^6

This latter finding is related to an important question for modeling the bias – if people ignore the lower digits, what do they replace them with? Possibilities raised in the literature are complete drop-off (replacing with 00, e.g. Gedenk and Sattler (1999)), statistical inattention (i.e. replace with 00, but only with some probability intrapersonally, e.g. DellaVigna (2009)), or replacing the lower digits with some statistic such as the mean or modal price ending (e.g. Basu (2006); Chen et al. (2010); Gabaix (2014)). Unfortunately, current evidence cannot convincingly separate between these alternatives. One alternative that appears to be less likely is replacement with the modal price ending, which is 99. This explanation is challenged by the finding of 99 being the least likely to be recalled correctly, and by anecdotal quotes of people referring to “2.99 feels a lot less than 3.00”.

3.3 Price distribution is discrete, many prices end with 9

The first evidence for the bias is the extent of sorting itself. These are the revealed preferences of the pricing firms, showing that they choose to price at 9 ending prices very frequently. I calculated the shares using various data sets. Of 375 million observations of weekly prices from 250 stores of one national US retailer I find that 20% of prices end with 99 (if restricted to regular, not on-sale, prices the number increases to 27%) and 61% (85% of regular prices) end with 9. For 12 products (from DellaVigna and Gentzkow (2017)) across 70 chains and 11,000 stores, 30%-35% of prices end with 99, and 70%-86% end with 9. In Israeli data regarding 20 products across 110 stores of different supermarket chains, I find that 45% end with 99. In the entire micro data of prices collected by the Israeli Central Bureau of Statistics (for the creation of the Consumer Price Index), roughly 50% of prices end with 99. In other papers, similar numbers arise. Levy et...
al. (2011) find 15% of 99s in another US retailer (though measurement issues probably mean this is a lower bound), and Conlon and Rao (2016) find 91% of 99s for Spirit products using Nielsen data. That is, across settings, retailers, countries, and time, firms behave as if 9 and 99 ending prices are superior to other price endings.

3.4 Evidence from Pricing reforms

Another piece of supportive evidence comes from countries who faced a sudden change in allowed prices. Notably, many countries that joined the Euro zone had stopped using their denomination in favor of the Euro. The currency change naturally caused a shift in the price distribution. As some papers document (Aalto-Setala (2005); El Sehity et al. (2005)), 9 ending prices were over-represented in national currency before the introduction of the Euro (in El Sehity et al. (2005) about 24% across countries), and a few years after the introduction 9s were common again (19%). However, one might argue that joining the Euro zone was much more than a mere currency change, and that many shocks occurred simultaneously. A cleaner reform happened in Israel at the end of 2013 (Ater and Gerlitz (2017); Haaretz (2015)). There, the only change was that the government banned pricing in non-0-ending prices (e.g., 2.99 was banned and so were 2.95 or 2.93, but 2.90 or 3.00 were allowed). As Shlain (2018) shows, the shares of 99-ending prices were hovering around 45% before the reform, and about a year after the reform the share of 90-ending prices were about 50%-60%. This has happened even though many chains, during the first months after the reform, were jolted out of the 99 equilibrium. These firms experimented with 00-ending prices, but then converted to 90-ending prices, thus providing additional indirect support for the existence of a left-digit bias.

3.5 alternative explanations

What else, other than left-digit bias can explain the discreteness and excess mass at 9 ending prices? One possible, yet implausible, explanation is that the products unit costs are also discrete in such a way that they happen to align to be translated to 9 ending prices. Using the retailer data I can reject that explanation. Figure 2 shows the price-ending distributions of 238 million product-week-store observations, for the “regular price” that consumers face, and the wholesale price the retailer pays for the products. The price-ending distribution of regular prices is discrete, in the sense that 25% of prices end with 99 (which is one of 100 possible price endings), and about 85% of prices end with X9 (one of 10 possible price endings). In comparison, the wholesale prices appear to be roughly uniform with no preference for 99 or 9 ending prices. So, we can rule out the case that the cost distribution is as discrete as the price distribution. More importantly, it is also clear that the 9-ending pricing behavior is only arising when facing customers, yet the manufacturers selling their goods to retailers do not price in a similar fashion. That is, the patterns are linked to ordinary consumers but not when the consumer is a corporation.

Importantly, there is also an alternative notion of the effect of 9-ending prices as being uniquely beneficial. That is, demand is downward sloping, but is locally increasing (sometimes referred to as a “blip”) at 9-ending prices. Some papers do find support for that idea (notably Anderson and Simester (2003)), but the econometric specification there can not separate between this explanation and the simpler left-digit bias idea. Further, if that is the case, missing prices around 99 should be symmetric above and below 99. As already evident from 2 this is not the case.
Another possibility is that 9-ending prices are chosen to simplify the pricing problem of retailers. This is the spirit of the price discreteness explanation of Matejka (2015), showing that seller’s limited attention optimally leads to discrete prices, and also a limited-time “sale price” behavior. While the model is useful for explaining sale prices versus regular prices, and generate discreteness in those, the prices themselves are a function of the information process – they do not end with 9. In other words, to explain 9 ending prices, the Matejka model needs extra structure.

Other explanations differ in the underlying mechanisms, but they lead to the reduced form left-digit bias discussed above. For example, one possible explanation is that prices are discrete due to rational limited attention (e.g. Basu (2006); Chen et al. (2010); Gabaix (2014)). Though they have other predictions, to the extent that they lead to 9 ending prices, these models are equivalent in the observable price distribution to other models. Importantly, for 9 ending prices to appear repeatedly across domains, there must by something about them beyond the model. In the models of Chen et al. (2010) and Gabaix (2014) there is indeed nothing special about 9 and this is not a natural prediction of the models.

4 Data

This project relies on various data sources, aiming to provide a thorough investigation of the bias. Each data set has its own benefits and shortcomings. Together, they tell a consistent story.

4.1 Retailer

I got access to data from a US national grocer (used before in other papers such as Eichenbaum et al. (2011); Gopinath et al. (2011)). Each observation is a product (UPC) in a store in a week. As shown in column (4) of Table 1, there are about 74000 distinct products, in 250 stores, and 177 weeks (2004 to mid-2007). For each observation I observe 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 was priced at its list/regular price), and other variables (such as two cost measures - wholesale prices and adjusted gross profits). The price is the division of net revenue by units sold, and the regular price is the division of gross revenue by units sold. To get instruments for prices, I am using leave-out-prices (in the spirit of Hausman et al. (1994)) for the same product at the same period in other stores. To make these prices to more likely capture supply shocks rather than local demand shocks, I use stores in other regions, but within 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.

An important issue with such scanner data is that if a product did not sell in a certain week, it might be because it was out of stock, or because it was overpriced. Prices are not observed if the product did not sell at all, leading to a potential underestimation of price elasticity. To mitigate this issue, 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 10000 observations. To get some level of minimal price-ending and dollar variation, I am selecting on 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 (5) of Table 1. 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). For that reason I exclude these categories among others.

I exclude the following: Delicatessen, Food Service, Fresh Produce, Meat, Seafood, and Alcoholic Beverages.

7
As shown in Table 1, this procedure leaves me with 1,710 products and 51.3 Million observations. Since these are the most available products, while they are merely 2.3% of all possible products, they capture 19.2% of all revenue in the data.

For my purposes the exact price paid is crucial. However, the inferred price (net revenue divided by units sold) does not necessarily describe an actual price paid in store. This happens when some consumers paid with coupons, if some had club membership specific discounts, or if some products were under non-linear pricing (such as “buy 1 item get the second for 50%”\(^8\)). This is a known issue in the literature (e.g., Einav et al. (2010); Eichenbaum et al. (2014)). However, the retailer is better than Nielsen RMS in the sense that there is no rounding, e.g., a price in the data might be 3.044. In such cases, when the price is not “to the cent” it is safe to infer that it is 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 46.1 Million (column (6) of Table 1).

Some statistics about the price endings\(^9\) are shown in Panel B of Table 1. As we move from all observations to the “real prices”, the share of 9-ending prices increase, meaning that other price endings are more likely to be the result of averaging of multiple prices. Notice also the differences in the share of 0 ending prices for all prices (column (6)) versus only regular prices (column (7)). Meaning, 0-ending prices are strongly positively correlated with being on-sale.

If 9-ending prices are mostly not on-sale, while 0-ending prices (and others) are on-sale, this will lead to spurious higher demand for 0-ending prices. If 1-ending prices are mostly fictitious, as result of mixing sale prices with regular prices, they will also exhibit high demand that will be wrongly attributed to specific prices. Figure 3 shows the empirical probability of an observation with a rounded price to be a real price (left panel) and on-sale (right panel) conditional on the last digit of the price. Indeed, almost all (98%) 9-ending prices represent prices actually paid in the store, but only 9.9% (!) of them are on-sale, compared to 0-ending prices, that are also mostly real (92%), but 91% of them are on-sale prices. 5 is an in-between case - mostly real prices, and less likely to be on-sale price, and other prices are rare, but likely result from averaging various prices, and likely on-sale.

4.2 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 records 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 a subset selected by DellaVigna and Gentzkow (2017), of 12 items (termed “high-end” in their paper) in food stores, since their main 10 items sample includes mostly products with average price below $1, and therefore absent dollar-crossing variation. In these 12 items also included is one of their main 10 UPCS, orange juice, with an average price of $3.54\(^10\). The summary statistics are described in column (1) of Table 1.

---

8 Another potential source of multiple prices, price changing 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.

9 Here I rounded the price to be to-the-cent, as happens for Nielsen data.

10 This sample will be expanded in the future.
The scanner data is similar to the retailer, with the key advantage being that it captures a larger section of the retail market, and 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. First, prices are always “to the cent”, not allowing to infer if an observation represents a fictitious price. Second, Nielsen measurement weeks are not always aligned with a retailer’s price changing frequency. 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 price mixing, it is unlikely to be identical for two consecutive weeks. 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. Figure 4 is 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 1 week spell, and that share jumps to 95% already for 2 week spells, and close to 100% for 3 weeks and longer. The yellow triangles show that share of observations that are regular prices (rather than 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.

5 Demand Estimation

In this section I will provide reduced form support for the bias in the form of discontinuous demand slopes, and then proceed to estimate the bias parameter as prescribed by the model. Finally, given estimates at the product level, I will investigate how do estimates change with respect to underlying covariates of the product and its clientele.

\[11^{th}\] Mostly true for a popular product purchased in multiple units each week, otherwise might be likely.
5.1 Reduced Form Evidence

Figure 6(a) shows simulated demand curves in log-log space, with elasticity of -4 and bias of 0.15 (and no added error) on the left panel, and a no-bias case on the right. The dashed line is the fitted line from a log(quantity) on log(price) regression. In turn, Figure 6(b) shows the residuals by price from these regressions. This seesaw pattern of residuals for the bias case, with discrete drops at dollar thresholds is a manifestation of the bias, doing the same exercise with real data. Note that if there is no bias, as in the right panel of figure 6(b) is a flat line at 0. If the model is misspecified, it will exhibit some other shape.

As evident from Figure 6, the residuals are better for an eye-test, since differences are more pronounced. Focusing on the residuals also has the advantage of allowing for aggregation of different elasticities (different slopes).

[Figure 6 about here.]

Starting with an example of a single good, here 10oz Lorna Doone Cookies, Figure 7 shows the actual data equivalents of Figure 6 in the Nielsen data. The left panel is the demand curve, aggregating over 12 retailers for which estimated elasticities are similar (-1.3 to -1.4). I aggregate log-quantity sold, residualized from time and store fixed effects (details below), into 10 cent bins. The size of each circle represents the number of observations. Different colors represent different dollar digits. The patterns are supporting the model, as differences in demand between same dollar-digit prices are smaller than the differences when the dollar-digit changes. Panel (b) shows the results of the residuals, now for all 52 retailers (hence with different price elasticities). Looking at the residuals, the patterns are starker, and the seesaw shape is clear. The patterns imply that changing the price, holding everything else constant, from $3.99 to $4.00 is linked to an average drop in demand of 0.085 log points. Visually, the patterns are aligned with the model (as in Figure 6), but this is just an example for a single product.

[Figure 7 about here.]

Above, and in the following, the specification used is

\[
\log Q_{ist} = \epsilon_{i,\text{chain}(s)} \log P_{ist} + \beta_{is} + \gamma_{i,\text{year}(t)} + \delta_{i,\text{month}(t)} + \mu_{i,\text{spell.length}(ist)} + \nu_{i,\text{cents.digit}(ist)} + e_{ist}
\]  

(6)

Where \( i \) is product, \( s \) is store, and \( t \) is period. \( \epsilon \) captures the product-firm elasticity; \( \beta \) is a store shifter (market size); \( \gamma \) is a year first effect capturing shifts in demand; \( \delta \) is month-of-year fixed effect capturing product specific seasonality; \( \mu \) is a spell-length fixed effect for spells of length 2, 3, 4, 5 weeks, and more than 6 weeks, capturing on-sale propensity (motivated by Figure 4); and \( \nu \) is a fixed effect for the cents digit being 9, 0, 5, or other, capturing different promotional strategies (demand shifters) that are correlated with specific price endings\(^\text{12}\).

There is always a worry that can not be ruled out of elasticity bias driven by local demand shocks affecting prices and quantity sold, manifesting itself as omitted variable bias. First, much of the demand variation is absorbed by the fixed effects, capturing seasonality, permanent store differences, and sale strategies. As \( \text{Rossi (2014)} \) argues, long panels with multiple fixed effects might not be subject to substantial bias. Relatively, DellaVigna and Gentzkow (2017) show that firms use uniform prices across stores, hence implying that price variation is not driven by local shocks. Indeed, I will be using in many of the specifications here just an OLS (or a non-instrumented NLLS). But, to mitigate worries of endogeneity of the price, I instrument

\(^{12}\)Note that since the bias is estimated from discontinuities across dollar digits and not dime digits, these fixed effects do not affect the findings.
for $\log P_{ist}$ with $\overline{\log P_{ist}}$ which is the average log-price of product $i$ in period $t$ in other stores in different counties of the same chain $d(s)$. The assumption here is that this is capturing changes in price that are common across stores and hence are unlikely to be driven by local demand shocks. These instruments can not fully resolve the issue of endogeneity, since it might be that there are more aggregate demand shocks that affect prices across stores in a certain week.

Since I allow for heterogeneous elasticities at the product-chain level, aggregating multiple products and chains can not be displayed as a single residualized demand curve, so I defer to the residuals plots, averaging the residuals first within product as described above (meaning effectively by chain size and chain availability), and then averaging again equally across products. Figure 8 shows the aggregated residuals for 12 products from DellaVigna and Gentzkow (2017).

Figure 9 is very similar, but from a larger set of products using only the retailer data. The estimated equation in that case is slightly different:

$$\log Q_{ist} = \epsilon_i \log P_{ist} + \log P_{c(i)st} + \beta_is + \gamma_iyear(t) + \delta_i,month(t) + \mu'_i, on, sale_{ist} + \nu_i, cents, digit_{ist} + \epsilon_{ist}$$  \hspace{1cm} (7)

where there are a few differences from equation 6. First, there is a single product-level elasticity. Second, instead of using spells as proxies for on-sale propensity I am using an indicator for whether a products 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). Third, I added another term capturing cross-elasticities, where $\log P_{c(i)st}$ is the average log-price of other products of the same product category $c(i)$ in the same store $s$ at the same period $t$. Last, I instrument for $\log P_{ist}$ with $\log P_{id(s)st}$, the average log-price of the same item $i$ in same period $t$ in stores in other cities but the same distribution center $d(s)$.

As described above, a further advantage of that data is that exclusion is based on if the implied weekly price is “to-the-cent” or not, thus keeping roughly 90% of observations rather than 73% as in Nielsen. The results for 1,710 products of regular-price availability of at least 50% and at least 10,000 observations are shown in Figure 9. In appendix figure 16 I show further robustness tests – OLS vs the IV is very similar, and so is a semi-elasticity specification. The specification that looks most aligned with the model is when I restrict attention to estimating the demand response to regular prices only. This finding means that the control of sale effects is not flexible enough, and that sales effect are more than a fixed increase in demand.

Both for Nielsen and Retailer datasets, and across specifications, patterns are consistent and very similar. They demonstrate a clear seesaw pattern of residuals, showing demand drops of a few percentage points for each dollar threshold. While this is supporting the existence and soundness of the model, more is needed in order to translate these patterns to parametric left-bias estimates.

[Figure 9 about here.]

### 5.2 Structural Estimation of $\theta$

I know turn to connect the reduced form evidence with the model, and to estimate $\theta$. 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 do consumers ignore. Second, since the model nests no-bias as a special case (with $\theta = 0$), finding positive bias levels in a consistent way provides further support for the model. Third,
structural estimation allows to describe the demand curve with a functional form, and to find analytical solutions for the firm pricing in response to it. And finally, structural estimation enables counterfactual exercises regarding firm profits under various scenarios.

I estimate the left-digit bias parameter at the product level, that is, estimate multiple $\theta_i$s. The reasons for that are that it is computationally infeasible to estimate a single parameter with the size of data and required number of fixed effects and controls; but more fundamentally, as will be shown below, there is substantial heterogeneity in the bias estimates which is strongly correlated with underlying characteristics of the different products and their clientele. Learning about this heterogeneity further inform us about the underlying mechanisms driving the bias in the field.

5.2.1 Sufficient Statistics Approach

The bias can be inferred from the demand drops at dollar thresholds. For example, say 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 in a no-bias world. That is, as if the perceived price just below $5.00 is $4.90 

\[
\text{In general, for each triplet of demand drop } \Delta \log Q, \text{ elasticity } \epsilon, \text{ and dollar threshold } d, \text{ the bias is:}
\]

\[
\theta = d \cdot \Delta \log Q / |\epsilon|
\]

This is the first estimation approach, where I estimate the demand drops for each product-dollar pair, and complement it with the elasticity from the IV specification in Equation 7. To get the demand drops I estimate for each product

\[
\log Q_{ist} = \sum_{d=\text{dollar-digit}} (\epsilon_{id} \log P_{idst} + \kappa_{id}) + \eta_{i,year(t)} + \beta_{it} + \gamma_{i,month(t)} + \delta_{i,month(t)} + \bar{\mu}_{i,\text{on-sale(ist)}} + v_{i,\text{last-digit}(ist)} + \epsilon_{ist}
\]

clustering standard errors by store, and then project the drops for each product at each dollar digit $x$ as

\[
\Delta \log Q_x = (\hat{\epsilon}_{x-1} \log(x) + \hat{\kappa}_{x-1}) - (\hat{\epsilon}_x \log(x) + \hat{\kappa}_x)
\]

where I calculate standard errors using the delta method. Appendix Figure [17] shows the estimated drops for all products, where I keep for each product the most precisely estimated drop in case there are more than one crossed dollar threshold.

The key advantage of this estimation technique is that even if pricing is simultaneous with demand shocks, to the extent that these shocks affect levels but not sorting into specific prices, the drops are still unbiased estimates. In a rough analogy to regression discontinuity design, as long as there is no sorting due to demand shocks (sorting due to supply shocks are fine and even required), the projected drops at dollar thresholds are unbiased. Unlike in RDD, there is no clear way to test for this restriction using density tests, mainly because we do expect that the bias leads to price-ending selection.

The CDF of all product level estimates in the retailer data is the black curve in Figure [10]. Most estimates are between 0 and 1, with a mean (median) of 0.251 (0.212).
5.2.2 Non-linear least squares

An alternative to the sufficient statistic approach is to estimate the model via non-linear least squares. Namely, instead of regressing log-quantity on log-price, to regress it on log-perceived-price, and estimate the elasticity and \( \theta \) simultaneously:

\[
\log Q_{ist} = \epsilon_i \log \left( (1 - \theta) P_{ist} + \theta [P_{ist}] \right) + \alpha_i \log P_c(i)_{ist} + \beta_is + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu'_{i,\text{on.sale}(ist)} + v_i,\text{cents.digit}(ist) + \epsilon_{ist}
\]

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 \( \theta \) of 0.21.\(^{14}\)

The identifying assumption being that variation in prices 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 red curve in Figure 10, which is very similar to the estimates coming from the drops. The mean bias is 0.254 and the median is 0.18.

5.2.3 Semi-elasticity version

A third method is to assume a log-linear structure instead of a log-log estimation. This is not driven directly by the model, but it has the key advantage of plugging in the perceived price and estimating it directly. That is, instead of the standard log-linear structure of \( \log Q \sim P \), I regress \( \log Q \) on \( P + [P] \). That is, I run at the product level

\[
\log Q_{ist} = \eta_i (1 - \theta_i) P_{ist} + \eta_i \theta_i [P_{ist}] + \alpha_i \log P_c(i)_{ist} + \beta_is + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu'_{i,\text{on.sale}(ist)} + v_i,\text{cents.digit}(ist) + \epsilon_{ist}
\]

where \( \eta \) 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 give the ratio of \( \frac{1 - \theta}{\theta} \). Therefore, within an OLS framework, I can estimate the bias directly. Reassuringly, the estimates from that approach are almost identical to the estimates from the NLLS approach, as can be seen in Figure 10. Apart from the 10% of cases in which the estimated NLLS elasticity is problematic (very inelastic, greater than -0.3), the estimates have a correlation of 0.86 and a rank correlation of 0.94. Their mean bias is 0.227 with a median of 0.166.

5.3 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, while only 29% if it is out of $125. 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.\(^{[forthcoming]}\)). In this left-digit bias setting, a manifestation of relative thinking is if the bias is higher for high ticket items, everything else

\(^{14}\)Estimation is very stable, and does not change if I use other values of \( \theta \).
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 9, 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. 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 11. 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 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.

6 Firm Pricing

Building on the demand side estimation corroborating the

\footnote{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.}
6.1 Qualitative Predictions

As analyzed in section 2, the model prescribes three key qualitative predictions for firms facing a demand structure such as the one estimated in section 5. The model predicts for optimizing firms (1) excess mass at 99, (2) coming from a region of missing prices with low price-endings, (3) which is increasing with the dollar digit.

**Excess mass and missing prices** Figure 12 shows the price-ending histograms for the regular prices in retailer data and the long price spells in the Nielsen data (at least 6 weeks). First, 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)\(^{16}\). This 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 no variation in the last digits, and hence the drops in demand at the dime thresholds can not be identified. However, going forward in conducting the firm response calibrations, I will do robustness tests with additional bias regarding the last digits (that is assume that with some probability $$\theta_2$$ consumers see the price as the dollar and dime digits and ignore the cents).

Figure 12 also show that for both the retailer and the Nielsen data, there are significantly less prices ending with anything lower than 19. This finding has not been described, predicted, or explored before.

[Figure 12 about here.]

**Next-lowest price by first digit** Figure 12 masks the heterogeneity of how price endings change when the dollar digit change. 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. Note that the share of 99-ending prices is not clearly predicted, since it relies on the distribution of costs.

To succinctly present it, I display the CDF of price endings separately for each dollar digit in Figure 13. Darker lines are higher dollar digits. Indeed, Figure 13 shows that the entire distribution of price endings is roughly shifting to the right, meaning that for higher dollar digits there are less low price endings and higher shares of 99s.

[Figure 13 about here.]

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

6.2 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 5? In the following section I present some predicted “as if” pricing by a firm facing uniform cost distribution and a single elasticity and bias level\(^{17}\) and compare those to the actual pricing behavior of the retailer.

\(^{16}\)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%).

\(^{17}\)In progress is full structural estimation, rather than the rough calibration presented here.
Recall from Proposition 2 that next-lowest prices are decreasing as demand is more elastic and as the bias is weaker. Therefore, to get conservative predictions, I take a low-end bias level, of 0.16, and the mean bias from the IV specification of -1.5. For these levels, the predictions are stark – next-lowest prices are 99, meaning that all prices should end at 99 regardless of the cost distribution, shown in panel (b) of Figure 14. The results still hold even if we allow for mild heterogeneity in elasticities or bias levels.

Keeping in mind that firms do respond to the bias, qualitatively, I probe what can explain the discrepancy between the actual pricing and predicted. The first simple alternative is that elasticities are estimated with bias, but in reality (known to the firms) demand is more elastic. Panel (c) of Figure 14 shows the CDFs of price endings under the assumption that the bias is 0.16, but demand is ten times more elastic at -15. Clearly this is not a reasonable level of demand elasticity, but even at these levels, missing prices and masses at 99 are too high compared to the actual behavior in panel (a). Another reason to doubt the differently perceived elasticity argument is that 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 is not the case for the bias parameter, since this paper is the first to quantify it (in the academic circles). Indeed, if assuming as in panel (d) of Figure 14 that the elasticity is correct but the bias is perceived to be much lower, then the patterns seem more aligned in the shares of 99 and missing prices.

In theory, the above exercise can be conducted much more literally. 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 locally constant price elasticity and some parametric cost distribution. This is of great interest methodologically, since it allows for price elasticity identification just from the price distribution even without knowledge of costs or quantities. Price data are becoming more and more available (e.g., Cavallo (forthcoming)), and this is a rare case in which a behavioral bias is allowing for estimation rather than ruining it. This is because this added left-digit bias parameter is creating discontinuities that can be exploited. In a sense, this is the dual of the notch estimation in the taxation literature (see Saez (2010); Kleven and Waseem (2013)). 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.

Instead, I will pursue an in-between approach making stronger assumptions about the cost distribution, coupled with random effects of the bias and elasticity, to conduct a full maximum likelihood estimation. This is still work in progress, to be added in the coming weeks.

7 Counterfactuals

At this point, we have a model and estimation of the demand and supply sides, allowing to ask different questions. First, how different are firm profits with versus without the bias? Second, if firms were to price according to the model, how higher were their profits versus a case in which they ignore it altogether? And finally, if firms price as if the bias is much lower, how much profits do they lose?

The results of this exercise are shown in Figure 15. Profits are shown against varying levels of bias, assuming a price elasticity of -1.5 and uniform cost distribution. I normalize profits to 1, shown by black circles, if the firm prices as if there is no bias. That is, according to the standard Lerner rule, where \( p = c + \frac{c}{1 + \epsilon} \). The yellow triangles show that the existence of the bias increase profits even if pricing does not change. This is not
extremely interesting since this is by assumption that demand is rotated upwards for all prices but round prices. For bias of 0.16-0.25 the addition to profits from this higher demand is 5-8 percentage points.

The blue squares are the added profits if the firm is optimally pricing according to Equation 3. For a bias of 0.16-0.25 this adds 1.5-4.5 percentage points of profits. Finally, the most interesting comparison is where do firms profits lie if the firms are perceiving the bias to be much smaller than it is, as calibrated in Section 6. This is represented by the green pluses, showing that most of the possible gains of optimal pricing are lost, and profits are better than pricing as if there is no bias at all by only 0.3-0.9 percentage points. That is, under-appreciation of the bias leads to significant profit losses of 1-3.5 percentage points.

Firms profits do not change due to large changes in the overall price level. As Appendix Figure 19 shows, and as can be inferred from the optimal pricing formula, the average price actually do not change much, and if anything is somewhat lower. The reason being that interior prices go up, but 99-ending prices mostly are coming from above. For less elastic demand, most of the prices shift down rather than up when there is a bias. But, the main effect on profits in the optimal versus naive pricing, is coming from not pricing at dominated regions. That is, not pricing below the next-lowest prices. 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.

[Figure 15 about here.]

A Math Appendix

A.1 Proof - Proposition

To see that equation desribes 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

\[
\frac{\partial \Pi}{\partial p} = 0 = Ae ((1 - \theta) p + \theta p_1)^{\epsilon - 1} (1 - \theta) (p - c) + A ((1 - \theta) p + \theta p_1)^{\epsilon - 1} (1 - \theta) (p - c) + \frac{A \epsilon}{1 - \theta} (1 + \epsilon)
\]

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

\[
p = \epsilon \frac{1}{1 + \epsilon} - \frac{p_1}{1 - \theta} \frac{\theta}{1 + \epsilon}
\]

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

18Since the differences between the bias and no-bias settings are by assumption positive, in Appendix Figure 19 I take the other extreme approach, and assume instead that when consumers are inattentive (i.e. with \( \theta \) probability), they replace the price with the ceiling instead of the floor. e.g., a $3-something price is perceived as $4.00. There, profits are of course lower with the bias versus without it. However, the effects of optimally responding to the bias and the underestimation of it are very similar. So the implications of underestimating the bias are, in that sense, robust.

19Whether consumers replace the price with the floor, ceiling, or any other arbitrary component does not affect this result.
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 \( q = p_1 - 0.01: \)

\[
D(p^*)(p^*-c) + D(p^*)(c-\bar{c}_q) = D(p^*)(p^*-\bar{c}_q) \\
< D(P)(P-\bar{c}_q) \\
= D(q)(q-\bar{c}_q) \\
= D(q)(q-c) + D(q)(c-\bar{c}_q) \\
\iff D(p^*)(p^*-c) < D(q)(q-c) + (D(q) - D(p^*)) (c-\bar{c}_q) \\
< D(q)(q-c)
\]

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 \( \theta \) low enough, such that there is not another point where demand is crossing the isoprofit.\(^{20} \)

### 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 \( \Pi(p^*,c) > D(q)(q-c) \). Define \( p' \) as the price that given demand \( D(q) \) results in the same profits as \( \Pi^* = \Pi(p^*,c) \). That is,

\[
p' = \frac{\Pi^*}{D(q)} + c \\
= \frac{\Pi^* - \Pi_q}{D(q)} + q + (c-\bar{c}_q) \\
\Rightarrow p' - q = \frac{\Pi^* - \Pi_q}{D(q)} + (c-\bar{c}_q) \\
> (c-\bar{c}_q) \left( 1 - \frac{D(P)}{D(q)} \right) > 0
\]

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

\(^{20}\text{i.e., } \theta \text{ needs to satisfy that } \Pi(q,\xi_\theta) \geq \Pi(q-1,\xi_\theta)\)
A.2 Proof - Proposition 2

Rewriting equation 5,
\[ P + \frac{(1 - \theta)P + \theta P_1}{\epsilon (1 - \theta)} - \frac{((1 - \theta)P + \theta P_1)^\epsilon P - ((1 - \theta)q + \theta q_1)^\epsilon q}{((1 - \theta)P + \theta P_1)^\epsilon - ((1 - \theta)q + \theta q_1)^\epsilon} = 0 \]

\[ P \left(1 + \frac{1}{\epsilon}\right) + \frac{TP_1}{\epsilon} \left(\frac{e P}{P + TP_1} + \ln \Lambda\right) + \frac{\partial P}{\partial \epsilon} \Lambda^e = \frac{\partial P}{\partial \epsilon} + \frac{\partial P}{\partial \epsilon} e + P - q \]

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

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

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

\[ (P - q) (q + T q_1)^\epsilon - (P + TP_1)^\epsilon + 1 \ln \left(\frac{P + TP_1}{q + T q_1}\right) > 0 \]

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

\[ (q + T q_1)^\epsilon - (\epsilon + 1) (P + TP_1)^\epsilon \ln \left(\frac{P + TP_1}{q + T q_1}\right) - (P + TP_1)^\epsilon > 0 \]

\[ \Leftrightarrow 1 > \Lambda^e (1 + (1 + e) \ln \Lambda) \]

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

where \( T \equiv \theta_1 - \theta \), the likelihood ratio of being inattentive and \( \Lambda \equiv \frac{P + TP_1}{q + T q_1} \).

A.2.1 The effect of elasticity on \( nI \)-price

First, note that

\[ \frac{\partial \Lambda}{\partial \epsilon} = \frac{\partial P}{\partial \epsilon} \frac{1}{q + T q_1} \]. Then, differentiating equation 9 with respect to \( \epsilon \),

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

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

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

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

\[ (P - q) (P + TP_1)^\epsilon - (P + T q_1)^\epsilon + 1 \ln \left(\frac{P + TP_1}{q + T q_1}\right) > 0 \]

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

\[ (q + T q_1)^\epsilon - (\epsilon + 1) (P + TP_1)^\epsilon \ln \left(\frac{P + TP_1}{q + T q_1}\right) - (P + TP_1)^\epsilon > 0 \]

\[ \Leftrightarrow 1 > \Lambda^e (1 + (1 + e) \ln \Lambda) \]

and note that \( \Lambda^e < 1 \), and that \( \ln \Lambda > 0 \) and \( \epsilon < -1 \) hence \( (1 + (1 + e) \ln \Lambda) < 1 \).
A.2.2 The effect of inattention on nl-price

How does higher inattention affect the next lowest price? Differentiate equation 9 with respect to $\theta$:

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

the term in parentheses on the left-hand side is negative, so the effect is determined by the opposite sign of

$$
\frac{P_1}{1 + \epsilon} \left( \frac{q + Tq_1}{p + TP_1} \right)^\epsilon - P_1 + \left( 1 - \frac{q_1}{q + Tq_1} \right) \frac{P + TP_1}{q + Tq_1} < 0.
$$

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

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

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

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

then it is left to show that

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

or

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

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.

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. Solving the discrete case is still underway.

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 $\theta$ 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 5.** There is no pure strategy equilibrium in this game

**Proof.** Assume all firms play a strategy \( p \). If \( p > 0 \) an \( e \) 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 6.** (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 \( \theta \) 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 \( \phi \) 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

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

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

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.\(^{21}\)

\[^{21}\text{If there is a mass point at 1, then it is profitable to shift it to just below 1 and gain } \theta \frac{1 - (1 - \phi)^N - (1 - \phi)^{N-1} \phi}{N\phi} = \theta \frac{1 - (1 - \phi)^{N-1}}{N\phi} > 0\]
Also, within each region there are no gaps. Assume there is, i.e. \( f(p) = 0 \) for all \( p \in \text{supp}(\overline{p}) \) where either \( \overline{p} < 1 \) or \( \overline{p} > 1 \). 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,

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

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

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,

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

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

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 \( \phi \), 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 \( \phi \) can solve that
equation. So, a necessary condition for such an equilibrium to exist is that

\[ 2 - N \frac{1 - \theta}{\theta} > 1 \]
\[ N \frac{1 - \theta}{\theta} < 1 \]
\[ N < \frac{\theta}{1 - \theta} \]
\[ \theta > \frac{N}{1 + N} \]

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 \( \theta > \frac{N}{1 + N} \). 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 \( \theta \) 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 Multiproduct Monopolist

D Appendix Figures

[Figure 16 about here.]  
[Figure 17 about here.]  
[Figure 18 about here.]  
[Figure 19 about here.]  

References


## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Illustration of the optimal price schedule</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>Price ending distributions for regular and wholesale prices</td>
<td>34</td>
</tr>
<tr>
<td>3</td>
<td>Real prices and on-sale by last digit of rounded weekly prices</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>Real prices and regular prices by length of price spell</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>Price ending shares and sample selection</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>Illustration: Demand and residuals with left-digit bias</td>
<td>38</td>
</tr>
<tr>
<td>7</td>
<td>A single product demand and residuals</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>Residuals of 12 products using Nielsen data</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>Residuals of 1710 products from a single retailer</td>
<td>41</td>
</tr>
<tr>
<td>10</td>
<td>CDF of left-digit bias estimates on Retailer data, three estimation methods.</td>
<td>42</td>
</tr>
<tr>
<td>11</td>
<td>Simple correlates of product and clientele characteristics and estimated bias</td>
<td>43</td>
</tr>
<tr>
<td>12</td>
<td>Price Ending Distributions</td>
<td>44</td>
</tr>
<tr>
<td>13</td>
<td>Price endings CDFs by dollar digit</td>
<td>45</td>
</tr>
<tr>
<td>14</td>
<td>Predicted and actual price-endings distributions</td>
<td>46</td>
</tr>
<tr>
<td>15</td>
<td>Counterfactual exercises</td>
<td>47</td>
</tr>
<tr>
<td>16</td>
<td>Residuals of 1710 products in Retailer data, different specifications</td>
<td>48</td>
</tr>
<tr>
<td>17</td>
<td>Estimated demand drops with standard errors</td>
<td>49</td>
</tr>
<tr>
<td>18</td>
<td>Counterfactual exercises</td>
<td>50</td>
</tr>
<tr>
<td>19</td>
<td>Relative price level by bias</td>
<td>51</td>
</tr>
</tbody>
</table>
Figure 1: Illustration of the optimal price schedule
Optimal price of a monopolist by cost with left-digit biased consumers. Gray line is the no-bias counterfactual. In purple are the prices that satisfy the interior solution, and in green 99-ending prices.
Figure 2: Price ending distributions for regular and wholesale prices
All retailer data from 238 million observations of products priced at their regular price. The top panel shows the price ending histogram of regular prices, which are the prices of a product if it is not under some transitory sale. All tall bars are at 9-ending prices. The bottom panel shows the price ending histogram of the same products, but for the wholesale prices the retailer paid to manufacturers.
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.

Figure 3: Real prices and on-sale by last digit of rounded weekly prices
Figure 4: Real prices and regular prices by length of price spell
Figure shows shares by price spell length for the Retailer data. An product-week observation is allocated to spell length $T$ if the price on that week is identical across $T$ consecutive weeks. The light blue line is the CDF of observations in the Retailer data. The black circles are the share of observations, conditional on spell length, that are “real” (i.e., prices ending to-the-cent). The yellow triangles are the share of observations, conditional on spell length, that are “regular” (i.e. net price equals regular price).
Figure 5: Price ending shares and sample selection
Figures show the shares of 9-ending prices, 0-ending prices and others in the Retailer data (left panel) and Nielsen data (right panel). The black bars are the data before cleaning (corresponding to columns (5) and (1) in Table 1). The yellow bars are the shares of “real” prices in the data (columns (6) and (2)) defined as to-the-cent prices in the Retailer data and at least 2 week price spells in Nielsen. The light blue bars are the shares for “regular” prices (columns (7) 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.
(a) Illustration: Demand with left-digit bias

(b) Illustration: Residuals from log-log regression with left-digit-biased demand

Figure 6: Illustration: Demand and residuals with left-digit bias
Actual demand minus predicted demand

(binned residuals from product–level regressions of $\log Q_{st} \sim \eta_c(s) \log P_{st} + \text{store} + \text{year} + \text{month}$ fixed effects, where $s$ is store, $c$ is chain, and 't' is date.

Chains with estimated elasticities of $-1.3$ to $-1.4$.

Subset of `real prices` including only price spells of length 2 weeks at least.

Residuals are binned to 10 cents bins, and weighted by price frequency. Product: cookies

(a) Demand, product: Lorna Doone Cookies

(b) Residuals, product: Lorna Doone Cookies

Figure 7: A single product demand and residuals.
Residuals from a regression of logQ_{ist} = \text{eta}_c(s)*logP_{st} + \text{spell.length}_{ist} + \text{last.digit}_{ist} + \text{store.ic}(s) + \text{year}_{it} + \text{month}_{it}

Regressions at the productXchain level (12 products, ~60 chains), where \( i \) is product, \( s \) is store and \( t \) time. Excluding 1 week spells, aiming to keep prices that are really the prices paid in store. Residuals are binned to 10 cents bins, and weighted by price frequency. 20.5M observations

Figure 8: Residuals of 12 products using Nielsen data
Actual demand minus fitted demand.

Residuals are binned to 10 cents bins, and weighted by price frequency. 46.6M observations.

These products and price spells capture at least 50% of total revenue.

Product selection: all products that have at least 50% availability, and 10,000 observations.

Using data on all non-averaged prices (50% of all observations) from a national retailer dataset.

Regression at the product level (1710 products, 268 stores, 177 weeks), where i is product, s is store, and t is time:

\[ \log Q_{ist} \sim \eta_i \log P_{ist} + \beta_i \log P_{-ist} + \text{last.digit}_{ist} + \text{sale}_{ist} + \text{store}_{is} + \text{year}_{it} + \text{month}_{it}, \]

where \( P_{ist} \) is instrumented by \( \log P_i(-s)t \) – prices of same item in stores in other cities.

Binned residuals from IV regressions of logQ_ist ~ eta_i*logP_ist + beta_i*logP_(−i)st + last.digit_ist + sale_ist + store_is + year_it + month_it,

Residuals are binned to 10 cents bins, and weighted by price frequency. 46.6M observations.

Figure 9: Residuals of 1710 products from a single retailer.
cdf of product-level left-digit bias estimates. Black curve is estimates from drops in demand and IV elasticity estimates. Red curve is estimates using non-linear least squares method. Blue curve is estimates from a log-linear specification.

Figure 10: CDF of left-digit bias estimates on Retailer data, three estimation methods.
Figure 11: Simple correlates of product and clientele characteristics and estimated bias

(a) Average product price

(b) Per-trip purchased units

(c) Household head education level

(d) Household annual income
(a) Retailer Data

![Retailer Data Graph]

Source: 1710 products, regular prices, retailer.

(b) Nielsen Data

![Nielsen Data Graph]

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

Figure 12: Price Ending Distributions
Figure 13: Price endings CDFs by dollar digit
Figure 14: Predicted and actual price-endings distributions
Figure 15: Counterfactual exercises

no bias cf = the standard case, with no-bias, and pricing according to a markup rule.
optimal pricing = pricing following the model, with higher markups for interior solutions and bunching at 99.
naive pricing = pricing as if there is no bias, while the bias is attenuating demand.
underestimating = pricing according to the model, but as if the bias is 100 times lower at 0.0016.
Figure 16: Residuals of 1710 products in Retailer data, different specifications

Figure shows the residuals as in Figure 9 for different specifications. Clockwise: Top left panel is “all”, an OLS without excluding the 10% of observations with non-real prices. Top center panel, “regular prices”, is for an OLS specification excluding all on-sale prices (keeping two thirds of the data). Top right panel, “semi elasticity iv” is similar to the main specification 7, only in a log-linear form, i.e., $\log Q \sim p$, instrumenting for the price with the price of the same item in other stores of the same distribution center in other cities. The bottom right panel, “to the cent iv lop”, is the main specification. Bottom center, “to the cent”, is the OLS version of the main specification. Finally, the bottom left panel, “semi elasticity iv w bias”, is an OLS modification of the semilog specification where the model perceived prices is introduced: $\log Q \sim p + |p| + FEs$, hence residualizing the bias.
product (sorted by drop size)

product-level drops in demand when changing dollar digits. Choosing one drop per product, 84\% of estimates are larger than zero (54\% significantly so at 5\% level); 16\% lower than zero (5\% significantly so). If taking all drops these numbers are 76\% (44\%) above zero, and 24\% (8\%) below zero.

Figure 17: Estimated demand drops with standard errors
Figure describes simulated profits for different bias levels and pricing strategies, with elasticity of $-1.5$.

- **no bias cf**: the standard case, with no bias, and pricing according to a markup rule.
- **optimal pricing**: pricing following the model, with higher markups for interior solutions and bunching at 99.
- **naive pricing**: pricing as if there is no bias, while the bias is attenuating demand.
- **underestimating**: pricing according to the model, but as if the bias is 100 times lower at 0.0016.

**Figure 18: Counterfactual exercises**
Figure 19: Relative price level by bias

average price by bias, conditional on elasticity and same underlying cost distribution

elasticity
-1.5
-2.5
-5
List of Tables

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summary Statistics and Data Selection</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>Correlations between product-level estimated bias and product characteristics</td>
<td>54</td>
</tr>
</tbody>
</table>
### Panel A: Data Description

<table>
<thead>
<tr>
<th></th>
<th>Full (1)</th>
<th>2+ weeks (2)</th>
<th>6+ weeks (3)</th>
<th>1710 UPCs (4)</th>
<th>Final (5)</th>
<th>Regular Prices (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations (M)</td>
<td>27.8</td>
<td>20.5</td>
<td>11.5</td>
<td>375</td>
<td>46.1</td>
<td>36.4</td>
</tr>
<tr>
<td>Products</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>74200</td>
<td>1710</td>
<td>1710</td>
</tr>
<tr>
<td>Stores</td>
<td>11000</td>
<td>11000</td>
<td>11000</td>
<td>250</td>
<td>248</td>
<td>248</td>
</tr>
<tr>
<td>Weeks</td>
<td>468</td>
<td>468</td>
<td>468</td>
<td>177</td>
<td>177</td>
<td>177</td>
</tr>
<tr>
<td>Chains</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Annual Revenue ($M)</td>
<td>97.9</td>
<td>57.2</td>
<td>28.4</td>
<td>2060</td>
<td>304</td>
<td>227</td>
</tr>
<tr>
<td>Annual Units Sold (M)</td>
<td>20.9</td>
<td>11.5</td>
<td>5.67</td>
<td>883</td>
<td>99.5</td>
<td>70.3</td>
</tr>
<tr>
<td>First Date</td>
<td>Jan-06</td>
<td>Jan-06</td>
<td>Jan-06</td>
<td>Jan-04</td>
<td>Jan-04</td>
<td>Jan-04</td>
</tr>
<tr>
<td>Last Date</td>
<td>Dec-14</td>
<td>Dec-14</td>
<td>Dec-14</td>
<td>May-07</td>
<td>May-07</td>
<td>May-07</td>
</tr>
</tbody>
</table>

### Panel B: Pricing Descriptives

<table>
<thead>
<tr>
<th></th>
<th>Full (1)</th>
<th>2+ weeks (2)</th>
<th>6+ weeks (3)</th>
<th>1710 UPCs (4)</th>
<th>Final (5)</th>
<th>Regular Prices (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Price</td>
<td>5.51</td>
<td>5.55</td>
<td>5.32</td>
<td>3.10</td>
<td>3.42</td>
<td>3.50</td>
</tr>
<tr>
<td>Price SD</td>
<td>2.81</td>
<td>2.78</td>
<td>2.57</td>
<td>1.97</td>
<td>1.63</td>
<td>1.64</td>
</tr>
<tr>
<td>Share on-sale</td>
<td>0.36</td>
<td>0.29</td>
<td>0.22</td>
<td>0.20</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>Share 99-ending</td>
<td>0.30</td>
<td>0.35</td>
<td>0.34</td>
<td>0.61</td>
<td>0.67</td>
<td>0.74</td>
</tr>
<tr>
<td>Share 9-ending (non-sale)</td>
<td>0.70</td>
<td>0.81</td>
<td>0.86</td>
<td>0.61</td>
<td>0.67</td>
<td>0.74</td>
</tr>
<tr>
<td>Share 0-ending</td>
<td>0.10</td>
<td>0.09</td>
<td>0.04</td>
<td>0.19</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Share 0-ending (non-sale)</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics and Data Selection
Dep: Left-digit bias

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(average price)</td>
<td>0.195***</td>
<td>0.141***</td>
<td>0.157***</td>
<td>0.210***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.030)</td>
<td>(0.036)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(per-trip units)</td>
<td>-0.489***</td>
<td>-0.332***</td>
<td>-0.396***</td>
<td>-0.220**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.050)</td>
<td>(0.076)</td>
<td>(0.089)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(income)</td>
<td>0.077</td>
<td>0.129</td>
<td>0.193*</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.104)</td>
<td>(0.113)</td>
<td>(0.114)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>education (z-score)</td>
<td>-0.011</td>
<td>-0.022**</td>
<td>-0.019*</td>
<td>-0.016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed Effects

<table>
<thead>
<tr>
<th></th>
<th>Category</th>
<th>SubCategory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>1,504</td>
<td>1,504</td>
</tr>
<tr>
<td>R²</td>
<td>0.094</td>
<td>0.076</td>
</tr>
</tbody>
</table>

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

Regressing product-level left-digit bias estimates on a product average price, and purchasing households per-trip units bought, income, and education (highest within 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’).