

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Essays in Behavioral Economics

Permalink

<https://escholarship.org/uc/item/9tt6j17q>

Author

Strulov-Shlain, Avner

Publication Date

2019

Peer reviewed|Thesis/dissertation

Essays in Behavioral Economics

by

Avner Strulov-Shlain

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Stefano DellaVigna, Chair
Professor Ulrike Malmendier
Professor Ben Handel
Professor Ned Augenblick

Spring 2019

Essays in Behavioral Economics

Copyright 2019
by
Avner Strulov-Shlain

Abstract

Essays in Behavioral Economics

by

Avner Strulov-Shlain

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Stefano DellaVigna, Chair

This dissertation studies empirically and theoretically how firms should and do respond to biased consumers. Specifically, I study the magnitude and implications of left-digit bias, the tendency of people to overweight left-most digits, on firms pricing behavior and profits.

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

In the second chapter, I ask to what extent do firms understand their consumers? When consumers are left-digit biased, demand is discontinuous at round numbers making them dominated prices. I present evidence that firms exploit left-digit bias in the long-run. They act as if they know this demand structure, using round numbers only for 1%-2% of posted prices. However, due to a slight reform in the set of admissible prices, supermarket chains in Israel fail to respond optimally - Round numbers post-reform account for 20% of prices, making pricing sub-optimal. After about a year firms re-optimize, and the shares of round numbers drop. Suggestive evidence show that having more spatial competitors is correlated with faster re-learning. Combining the well-understood bias, data, and reform, allows a unique window to look into firms sophistication level. I conclude that firms do respond to biases in the long-run, but do not have a model of these biases in mind and hence fail to respond in the short-run.

To Osnat

Contents

Contents	ii
List of Figures	iv
List of Tables	vi
1 More than a Penny's Worth: Left-Digit Bias and Firm Pricing	1
1.1 Introduction	1
1.2 Model	6
1.3 Data	11
1.4 Demand Estimation	14
1.5 Firm Pricing	23
1.6 Counterfactuals	28
1.7 Conclusions.....	31
2 Firm Sophistication: A Natural Experiment on Price Endings	47
2.1 Introduction	47
2.2 Literature.....	49
2.3 Model.....	50
2.4 Data and Setting.....	52
2.5 Empirical Analysis.....	54
2.6 Conclusions.....	60
Bibliography	71
A Appendix to Chapter 1	76
A.1 Proofs.....	76
A.2 Price Competition	81
A.3 Multi-product Monopolist	84
A.4 Nielsen Data	85
A.5 Bias Heterogeneity.....	86
A.6 Welfare	87

B Appendix to Chapter 2**107**

List of Figures

1.1	Demand curve of a single product, exhibiting drops at dollar digits	34
1.2	Illustration of the optimal price schedule	35
1.3	Illustration: Demand and residuals with and without left-digit bias	36
1.4	Sawtooth patterns of residualized demand by price.....	37
1.5	Histogram of product-level left-digit bias estimates from drops in demand at the dollar-digit	38
1.6	Pricing response: Price-endings distributions.....	39
1.7	Price endings: left-tail of CDF by dollar digit in retailer data	40
1.8	Estimated firm-perceived left-digit bias	41
1.9	Illustration of welfare effects.....	42
2.1	Minimal levels of left-digit bias $\theta_0(\epsilon)$ above which 00-ending is suboptimal ...	61
2.2	Shares of price endings over time.....	62
2.3	Shares of price endings post-reform by chain	63
2.4	Correlation between pre-shares of 99 and post-shares of 90 and 0. Each dot is a product in the ICC data.	64
2.5	Automaton tracking price endings path	65
2.6	Price ending paths of products who ended with 99 before the policy change ...	66
2.7	Spatial competition and shares of price ending.....	67
A.1	Data comparison, price ending shares, and sample selection	90
A.2	Data description: “real” and “on-sale” by last digit of observed price.....	91
A.3	Shares of real prices and on-sale prices by length of price spell	92
A.4	Retailer Robustness: Residualized demand of 12 products and 70 retailers using Nielsen data	93
A.5	Demand structure robustness: Residualized log-linear demand	94
A.6	Demand structure robustness: Residualized polynomial demand and rank fixed-effects.....	95
A.7	Sample robustness: Residualized demand for an alternative sample of products	96
A.8	CDF of left-digit bias estimates on Retailer data, three estimation methods. ..	97
A.9	Bias Heterogeneity: Correlates of product and clientele characteristics with left-digit bias	98

A.10 Simulation: Relative price level by bias	99
A.11 Illustration: pricing schedule of a 2-product monopolist.....	100
A.12 Minimum distance estimation: empirical and predicted moments	101
A.13 Estimated firm-perceived elasticity	102
B.1 Correlation between pre-shares of 99 and post-shares of 90 and 0. Each dot is a product in the CBS data.	108

List of Tables

1.1	Summary Statistics and Data Selection.....	43
1.2	Estimates of product-level drops and left-digit bias parameter	44
1.3	Estimated perceived parameters from firm pricing behavior	45
1.4	Profits lost when the firm underestimates left-digit bias.....	46
2.1	Summary Statistics of Prices Data	68
2.2	Correlation between 99-endings pre-reform and 00-ending prices in different periods.....	69
2.3	Price-ending effects	70
A.1	Summary Statistics and Data Selection, Nielsen Data.....	103
A.2	Correlations between product-level estimated bias and product characteristics.	104
A.3	Profits lost when the firm underestimates left-digit bias, alternative assumptions	105
A.4	Direction of welfare effects under various assumptions	106

Acknowledgments

This has been a wonderful journey thanks to the people who traveled along.

Stefano, you have actually been the advisor I wished I had. Your mix of enthusiasm and criticism is perfectly balanced. Always caring for me as a person and as a student. I remember sitting in 219B in my first year, and you were telling the class how some of the empirical papers in the class turned into oral projects and even dissertation chapters. Well, mine evolved slowly but surely, with your constant support, to a job market paper. You didn't give up on this project long after I already did. You have allowed me to explore, and then nudged me back exactly at the right time. Of course, you are an outstanding scholar - you don't need me to tell you that; but what I can attest to as your advisee and GSR for many years, is that you are also a truly good and kind person. I owe huge chunks of my success to you.

Ulrike - as Stefano admits, he sometimes have idiosyncratic tastes; with you I knew I could always get a second opinion. Your high standards and directness were always just the right thing and you pushed me forward again and again. Ben, you were supportive, kind, and frank, and were so generous with your time and efforts to support me before, during and after the market. Ganesh, you helped me tremendously with my introduction to marketing, a field I am ashamed to say, I wasn't too much aware of a short while ago. Finally, Ned, you were truly there in each and every meeting; completely engaging and thinking as hard as you can. I remember several meetings that changed the course of this project and jolted it ahead. Thank you all, you were my team, and I was lucky to have your support.

I had the privilege to be educated with many outstanding scholars who were generous with their advising time. Emmanuel, who guided a group of us through the third year and who taught me the power of good figures; David S, who opened my eyes to finance and who instilled security in my abilities; Dmitry, who was always engaging and pushed me to the frontier without cutting me any slack; Matthew, whose class I rushed to take in the first year before he left and am truly happy for doing so; and Shachar, who always provided me with a different perspective, creativity, and soberness. And also to many others who lent me their ears and taught me so much along the way: David A, David C, Philipp, Yoram H, Miguel, Pat, Supreet, Przemek, and Alan.

One of the best things during this PhD were the friends and colleagues I met. Mostly willing to listen to my complaints, stories, bad jokes, or unsolicited advice, and always willing to help. I was never alone thanks to my family, but I was also never professionally and socially alone thanks to you. Zarek, who joined the project fully at its previous incarnation, and was a true gentleman when it changed phases; Yotam, who gave me friendship, truth, and econometric support; My office mates and shark tankers along the years - Natalie, Julien, Waldo, Anne, Dmitri, Jen, Jon, Christina, Stephanie, Jimmy; My cohort friends who made this a true companionship - Juli, Nick, Daniel, Ale, Walker, Kevin, Rupal, Jonas, Eric, and Tamas; and those from different cohorts - David, Ferenc, Kati, Jordan, and Johannes; Finally, to Alex and Peter who I hope to keep working with down the road. I wish to add a special thanks to those who helped create and sustain the "Psych and Econ Therapy" group -

mostly Stefano and Alex. I was shamelessly pushing for it in my 4th year, since Stefano and Ulrike were away during my 3rd, where therapies are usually held. I think it had become an institution of high quality feedback and community building, helping to make Berkeley the best place to be a budding behavioral economist.

To my parents, Ayelet and Roni, who selflessly and consistently throughout my life allowed me to go down any educational route I chose. It is not easy to let your child go to a boarding school just because he wants to study with other nerds; it is not easy to send him with your daughter-in-law and only grandson across the ocean with blessings and hopes; and it is definitely not easy to support them when they wish to stay away a little longer.

Last but definitely not least, to my own little family. To Osnat - the love of my life; who brings the best out of me every day for years and years; who told me "we have no choice, you have to apply to a PhD in the US, I can't see you like that"; who sent me a link to an online article on a pricing reform in Israel "look at that, maybe you'll find it interesting"; who as cliché as it is, supported me through my PhD; who always always always has my back; who together brought and raises our two amazing children, Regev and Gevah. You three are my love, joy, and pride, you are what keeps my priorities straight and make me a happy man. I will do everything for you.

Thank you. Thank you all.

Chapter 1

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

1.1 Introduction

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

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

To answer these questions, I provide a portable model, examine whether it is supported by the data, and estimate its parameters. I integrate the structure of left-digit biased consumers of Lacetera, Pope, and Sydnor (2012) with a model of profit-maximizing firms to derive a set of predictions on pricing and profits. First, I provide reduced-form evidence for left-digit bias in consumer demand, and structurally estimate the degree of the bias on thousands of products in scanner data. I find that consumers are biased, to the extent of treating a 1 cent increase from a 99-ending price as if it were a 15-25 cent increase. Next, I estimate retailer pricing behavior. I find that firms pricing is aligned with the model’s main predictions,

¹See also Aalto-Setälä (2005), Anderson and Simester (2003), Anderson, Jaimovich, and Simester (2015), Ashton (2014), Ater and Gerlitz (2017), El Sehity, Hoelzl, and Kirchler (2005), Kalyanam and Shively (1998), Levy, Lee, Chen, Kaufmann, and Bergen (2011), Macé (2012), Schindler and Kibarian (1996), Snir, Levy, and Chen (2017), Strulov-Shlain (2019a)

but, they seem to underestimate the magnitude of the bias significantly. From the firm's perspective, I estimate that they act as if a 99-ending price is treated by consumers as being only 1.5-3 cents lower than the round price. Finally, I examine the effects on consumer surplus and deadweight loss and show that they are ambiguous in sign. However, because of the discrepancies between the estimated and firm perceived left-digit bias, I calculate economically significant implications on firm profits.

The first step in the paper is to create a framework to understand the phenomenon. As in Chetty, Looney, and Kroft (2009), DellaVigna (2009), Gabaix (2019), Lacetera et al. (2012), I model left-digit bias as if consumers ignore a θ share of a component of the price – the lower digits – and replace it with a focal price ending Δ . Focusing in this paper on single dollar-digit items, this makes a price p to be perceived as $\hat{p} = (1 - \theta)p + \theta([p] + \Delta)$ (a mix with weight θ on a the price with Δ cents component, and $1 - \theta$ on the exact price). For example, regardless of Δ , with a bias of $\theta = 0.15$, a consumer perceives the difference between \$4.00 and \$3.99 as being 15 cents, but between \$4.01 and \$4.00 as only 0.85 cent. The bias causes aggregate demand to drop when the first digit changes, e.g. from 3.99 to 4.00, and also dampens the price sensitivity to within-dollar price changes, e.g. from \$4.00 to \$4.01.

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

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

Figure 1.1 provides a motivating example of a demand curve for a single product. The horizontal axis is the price and vertical axis is log-quantity sold, demeaned from various controls. Details regarding data and estimation are described in the body of the paper. The figure shows that demand is flatter within dollars than across all prices (solid lines versus dashed line), with drops between dollars.

To present systematic evidence on a large sample of products, I residualize quantities purchased from the overall price sensitivity (e.g., with constant elasticity), and other product level controls such as seasonality, substitution with similar products, and promotions; then, I consider the *residuals by price* which are the unexplained component of demand. Relative to a smooth demand curve, residuals under left-digit bias have a sawtooth pattern - increasing

within each dollar digit and dropping sharply across the dollars. The main advantage of this procedure is that it allows to aggregate over different products (with different elasticities and price distribution).

Several empirical challenges require attention, of which I highlight two. First, the analysis hinges on precise measurement of prices, but prices in scanner data are the division of weekly revenue by weekly number of units. This effective weekly average price is the possibly a mix of different prices, for example if some consumers had a discount. To deal with it I exclude observations that are clearly mixes of different prices (for example if the price is not “to the cent”). Second, to estimate demand we need to capture price changes that are due to supply shocks. I control for most classic endogeneity concerns, and use instrumental variables inspired by Hausman (1996), with some advantages given by the richness of the data.

Figure 1.4 aggregates results from 1710 products and shows that the sawtooth patterns are clear. Demand indeed drops as the dollar digits change. The patterns are robust to the demand structure, product selection, retailers, and are cleanest for regular prices (i.e. when a product is not “on sale”).

These drops, together with the price level and price elasticity, can then be translated to estimate the bias parameter. The bias parameter describes by how much the price should change absent the bias to generate the same change in demand as from the discontinuous drop. Namely, consider a drop of 4.5 percentage points in demand crossing from \$4.99 to \$5.00. With an elasticity of -1.5, a change of 4.5% in demand can also be coming from a 3% change in price. Since 3% of \$5 are 15 cents, it implies a bias of 0.15, since for these parameters the drop in demand is the same as increasing the price from \$4.85 ($\approx (1 - 0.15) * 4.99 + 0.15 * 5$) to \$5.00 absent a bias.

I estimate the bias at the product level in 3 different methods, resulting in similar parameters with high pairwise rank correlation (0.91, 0.7). The average bias level is 0.16-0.24 depending on the estimation technique. I also estimate a specification with homogeneous left-digit bias and semi-elasticity, and estimate a bias parameter of $\theta = 0.16$ (0.01).

An important point is that whether perceived prices are rounded up or down does not affect these estimates, and is not important for the implications I study later in the paper. Meaning, I do not claim that total demand increases or decreases due to left-digit bias, nor do I test that claim.

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

The model predicts an optimal price given parameters of left-digit bias, elasticity, and cost. Discontinuities in demand lead to a pricing schedule by monopolistic firms that has distinct features (also for price-competing firms, in Appendix A.2, or multiple-products monopolist in Appendix A.3). Prices bunch at 99-ending prices, and there are well-defined regions of missing prices above the 99-ending prices (but not below). This latter prediction separates this model from the few other models suggested in the literature.

Qualitatively, firms pricing behavior follows the predictions and comparative statics of the model. First, retail prices commonly end with 99 (24%-34% of popular products). Second, they rarely end with a price-ending lower than 19². A third prediction of the model is that the range of missing prices is increasing with the dollar digits. This is also supported in the data. Low price endings are used less and less as the dollar digit increases.

However, quantitatively, there is a strong discrepancy between what the estimated bias imply for firms pricing patterns and what they actually do. Under the model, the estimated levels of left-digit bias (0.25 or even 0.16) and price elasticity imply that *all* prices should end with 99. If the firm believes the bias to be lower or demand more elastic, that can explain its pricing decisions.

Therefore, I estimate the firm's *perceived* left-digit bias and elasticity that rationalize their pricing behavior under the model. The estimation technique allows to estimate the perceived price elasticity solely from the price distribution, even absent knowledge of quantities or costs. The estimation technique relies on strong assumptions, and the estimates are noisy, but reasonable. Compared to demand-side estimates, I find perceived elasticity to be more elastic and the perceived left-digit bias to be much lower at 0.023, and even 0.005. I can reject a perceived bias larger than 0.043, which is lower than the demand-side estimated levels by a few factors. The estimated low perceived bias is driven by the excessive prevalence of low price endings.

Next, I investigate the welfare effects of the bias existence, and of the firm response to it. The bias might improve consumers surplus, because it pushes down some prices to be 99-ending, thus benefiting all consumers who now pay less for an item. While consumption distortion exists, it may not be the dominating force. Overall, welfare effects are ambiguous in sign due to the bias and to firms response to it.

Finally, given the demand side estimates, I ask what are the firm profits when it underestimates the bias. Namely, I compare counterfactual profits from ignoring the bias or underestimating it to a benchmark of full optimization. Compared to underestimating the bias, gross profits for optimizing firms are higher by 1-3 percents. A conservative estimate of the effect - considering only regular (non-promotional) sales - is that annual revenues for the chain are about \$60m lower due to underestimation of the bias.³

How can firms respond to the bias but at the same time underestimate it? First, this is not the first paper to document firms making consistent deviations from profit maximization (e.g. Bloom and Van Reenen (2007), Goldfarb and Xiao (2011)). Yet, a specific explanation for the partial correction seems plausible. Changing a naive \$2.00 to \$1.99, is intuitive and brings large gains, but a full correction, such as changing a naive \$2.40 to an optimal \$1.99, is counter-intuitive and less impactful on profits (and hence also harder to pin down with noisy data). In the final section I discuss alternative explanations, such as the model being

²When they do, they are almost always on-sale prices, and likely an artifact of specific promotion strategies (such as "2 for \$6")

³The result holds under various scenarios allowing for a multi-product price setting, or for heterogeneity in left-digit bias, perceived bias, or price elasticity. The losses are of the same magnitude whether consumers adjust prices downward or upward.

wrong or the plausible case in which consumers' bias level responds to the pricing behavior of the firm. I argue that these can not be ruled out, but seem unlikely.

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

Of course, the prevalence of 9 ending prices is not new, and has been extensively explored. Some papers draw attention to documenting the fact that many prices end with 9 (Hackl, Kummer, and Winter-Ebmer (2014), Levy et al. (2011)) or revert to 9 following currency changes (Aalto-Setälä (2005), El Sehity et al. (2005), Strulov-Shlain (2019a)), as a “revealed preference” argument for its benefit. Experimental papers explore underlying mechanisms (e.g. Carver and Padgett (2012)⁴). Fewer papers examine the effects of 9-ending prices on demand, either in field experiments (e.g. Anderson and Simester (2003), Ashton (2014)⁵), or in observational data (e.g. Stiving and Winer (1997)⁶). These papers find that 9-ending prices increase demand, with some mixed results. This paper adds to this literature by providing visual reduced-form evidence for the bias using big data, and a structural estimation and interpretation of the findings allowing to import the bias into the firm's problem⁷. This is also the first paper that quantifies firm response to the bias.

On the formal theory side, few models relate directly to 9-ending prices (Basu (1997, 2006), Stiving (2000)), and some bounded-rationality models generate discrete price-setting behavior (Chen, Iyer, and Pazgal (2010), Gabaix (2014), Matějka (2015)). This paper provides a portable model that is straightforward to take to the data, connecting both the consumer and firm sides. A predecessor smooth model of left-digit bias (DellaVigna (2009)) was estimated in various economic settings, but not on prices (Busse, Lacetera, Pope, Silva-Risso, and Sydnor (2013), Lacetera et al. (2012)). This difference is crucial, since prices are manipulateable⁸.

The paper relates to the literature in behavioral IO in two senses. First, by studying a market where firms respond to biased consumers (See Heidhues and Köszegi (2018) for review), and second by studying a case where the firms are also “behavioral”, in the sense

⁴Bizer and Schindler (2005), Carver and Padgett (2012), Schindler and Chandrashekar (2004), Schindler and Kirby (1997), Schindler and Wiman (1989), Snir et al. (2017), Thomas and Morwitz (2005). A summary table is available in Carver and Padgett (2012)

⁵See Anderson and Simester (2003), Ashton (2014), Bray and Harris (2006), Dalrymple and Haines Jr (1970), Dube, Manning, and Naidu (2017), Ginzberg (1936), Schindler and Kibarian (1996).

⁶See Blattberg and Wisniewski (1988), Hackl et al. (2014), Jiang (2018), Kalyanam and Shively (1998), Macé (2012), Stiving and Winer (1997). Observational data papers usually find higher demand for 99 or 9-ending prices by including a dummy for these. Surprisingly, these papers rarely discuss the issue of price averaging that is common across these data.

⁷A few papers document that prices tend to end with 9, or 99, and ask questions on the effect of this phenomenon on price stickiness (Anderson et al. (2015), Ater and Gerlitz (2017), Levy et al. (2011)) and the pass-through of taxes (Conlon and Rao (2016)). This paper adds to this literature by micro-founding this stickiness as an optimal pricing response of firms.

⁸Indeed, two papers that examine left-digit bias in house prices, find extensive sorting of prices, but try to argue that the price ending is chosen as-good-as random, which is clearly not the case in retail pricing (Chava and Yao (2017), Repetto and Solis (2018)).

that they are not profit-maximizing in a systematic way⁹. A companion paper, Strulov-Shlain (2019a), exploits a natural experiment in Israel and finds that firms do respond to left-digit bias in the long-run but in a way that is consistent with relearning a rule-of-thumb behavior rather than understanding the underlying model.

Finally, the paper relates to the vast literature of product demand estimation (see a review by Nevo (2011)), by contributing a novel method to estimate demand solely from the price distribution. This is a rare case in which a behavioral bias allows identification rather than make it impossible (see review by DellaVigna (2018)).

1.2 Model

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

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

Another popular class of models is of inattention to certain components of a price (such as sale taxes in Chetty et al. (2009) or shipping fees in Hossain and Morgan (2006)). The model was later translated to inattention to lower-digits by Lacetera et al. (2012) and estimated as the effect of odometer readings on prices in the used cars market. I use a generalization of this model, in which with probability θ a consumer ignores the right-most digits and replace them with a focal price endings (this is the generalization, in other papers the focal price ending is 0). In Lacetera et al. (2012), that number is the odometer of the car, and in this paper it is the price of a product.

Model The model consists of left-digit-biased consumers and monopolistic firms¹⁰. Consumers choose a quantity to purchase to maximize utility, but may misperceive the price.

⁹See Cho and Rust (2008, 2010), DellaVigna and Gentzkow (2017), Goldfarb and Xiao (2011), Hanna, Mullainathan, and Schwartzstein (2014)

¹⁰In appendix A.2 I solve a case of firms pricing under price-competition with left-digit biased consumers. The case of multi-product monopolist is in appendix A.3.

Firms price to maximize profits, according to their beliefs of consumer misperception. Misperception is a primitive of consumer behavior, and is not affected by firm behavior. In the last section of the paper I will discuss the plausibility and importance of this assumption.

Consider a product whose price is p . Assume that a consumer perceives the price as $\hat{p} = \hat{p}(p; \theta)$, a distortion function that causes the consumer to perceive the price as possibly different than it actually is. Given this paper's empirical context, the $\hat{p}$ function is such that consumers observe the left-most digits in full, but sometimes replace the lower digits with a focal price ending. Specifically, a left-digit bias parameter $\theta \in [0, 1]$ has two interpretations¹¹. One interpretation is that it describes that each consumer is putting a relative weight θ on the lower digits of the price, while a second alternative is that θ share of consumers always replace the lower digits with an arbitrary price ending. Though different in interpretation, these two modes lead to similar qualitative predictions (see Appendix A.2 for when they differ). The former interpretation is a bit simpler to solve algebraically, and also captures the observation that “4.99 feels lower than 5.00” at the person-specific level. This is the model of choice. That is, from now on assume that

$$\hat{p} = \hat{p}(p; \theta) = (1 - \theta)p + \theta(\lfloor p \rfloor + \Delta) \quad (1.1)$$

The focal price ending, Δ , is important for welfare and total profits (do consumers over- or under-consume?), but can not be identified in the data. In addition, it is less consequential for the impact of misoptimization by the firm. I highlight when this is consequential. Therefore, for simplicity and unless otherwise specified, assume that $\Delta = 0$.

If the agent has utility of the form $U(y, q) = y + \frac{(\frac{q}{A})^{1+\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}}$ where y is the residual income, q is the quantity purchased of a good, and A is translating quantity to numeraire value, then overall demand will be of the form¹²

$$D(p; \theta) = A\hat{p}^\epsilon = A((1 - \theta)p + \theta\lfloor p \rfloor)^\epsilon \quad (1.2)$$

To gain some intuition, the impact of the θ left-digit bias parameter can be understood by observing the effects of prices on demand: First, the penalty of a 0.01 price increase is large if the left-most digit changes (i.e. 2.99 to 3.00); Second, the effect of a 0.01 price increase on demand is lower than in the standard model if only one digit changes (i.e. 2.95 to 2.96).

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

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

¹¹In other papers, θ is termed “inattention”. However, other forces might lead to the same behavior, such as imperfect price recall, price categorization, or a tendency to choose round numbers for internal values or reference-prices. Therefore, I use the term left-digit bias which captures all of these, and remain agnostic about the behavioral mechanism.

¹²Conversely, if there is type mixture as described above, then following the same formulation will lead to demand of the form $D(p; \theta) = A[(1 - \theta)p^\epsilon + \theta\lfloor p \rfloor^\epsilon]$.

i.e., demand is driven by the *perceived* price while per unit profits are governed by the *true* price. This is a common feature of models with inattention (e.g. Farhi and Gabaix (2015), Gabaix (2019)), where there is a discrepancy between true utility and maximized utility on the consumer side, causing distortions.

In section 1.5 I consider a specific form of misoptimization by the firm. The monopolist understands the model and knows the true structural elasticity, but might be wrong about the level of left-digit bias, so will maximize profits given their belief. i.e., it maximizes $\Pi = D(p; \hat{\theta})(p - c)$, where $\hat{\theta}$ might be different than θ . For now, assume that $\hat{\theta} = \theta$.

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

With left-digit bias, we get a discontinuity in demand when changing p_1 since in the limit from below, demand is $A(p_1 - \theta)^\epsilon$ versus $A(p_1)^\epsilon$. To simplify the solution, assume that prices can be chosen from all real numbers between a natural number and a .99 ending number - $\bigcup_{q_1} [q_1, q] \subset \mathbb{R}_+$ where $q_1 \in \mathbb{N}$ and $q = q_1 + 0.9 + 0.09$. Proofs to the following propositions and corollaries appear in Appendix A.1.

Solving for optimal pricing yields the following: Optimal pricing formula, for small θ .

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

$$p(c; \theta, \epsilon) = \begin{cases} q & \text{if } c \in [\underline{c}_q, \bar{c}_q] \\ \left(c - \frac{\theta}{1-\theta} \cdot \frac{q_1+1}{\epsilon} \right) \frac{\epsilon}{1+\epsilon} & \text{if } c \in [\bar{c}_q, \underline{c}_{q+1}] \end{cases} \quad (1.3)$$

Where $\underline{c}_q$ and $\underline{c}_{q+1}$ are defined above, and $\bar{c}_q$ is defined below with an implicit equation as the minimal cost for which it is profitable to price strictly above q . To make Equation 1.3 clearer I elaborate on the three threshold costs $\underline{c}_q$, $\underline{c}_{q+1}$, and $\bar{c}_q$. The former two, $\underline{c}_q$ and $\underline{c}_{q+1}$, are the costs for which the monopolist's profit is maximized as an *interior* solution at q and $q+1$ respectively. i.e., where the first order condition is satisfied at q and $q+1$. The image of costs in $[\underline{c}_q, \underline{c}_{q+1}]$ is hence prices in $[q, q+1]$ which allows to use q and q_1 in the two cases in Equation 1.3. The third threshold, $\bar{c}_q$, is the cost for which prices will switch between q to a point in the segment $[q_1+1, q+1]$. At $\bar{c}_q$ two conditions are met: (1) the profit is maximized (with an internal solution) on some price $P_q \in [q_1+1, q+1] > q$; and (2) profits are equal at that price, P_q , and at q^{13} . For ease of notation, I omit q from P_q in the following:

$$\bar{c}_q = \begin{cases} P + \frac{(1-\theta)P + \theta P_1}{\epsilon(1-\theta)} \\ \frac{D_P P - D_q q}{D_P - D_q} \end{cases} \quad (1.4)$$

¹³For very low θ , there is no gap, and $P_q = q_1 + 1$.

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

$$P + \frac{(1 - \theta)P + \theta P_1}{\epsilon(1 - \theta)} - \frac{((1 - \theta)P + \theta P_1)^\epsilon P - ((1 - \theta)q + \theta q_1)^\epsilon q}{((1 - \theta)P + \theta P_1)^\epsilon - ((1 - \theta)q + \theta q_1)^\epsilon} = 0 \quad (1.5)$$

This equation is not analytically solvable, but can be easily solved numerically. Given P , we get $\bar{c}_q$ from either part of Equation 1.4. It means that no prices are set between $q_1 + 1$ and P .

To gain some intuition regarding Equation 1.3 note that this pricing behavior is different than a no-bias world in two ways. First, the top case shows that prices bunch at 99, and second, the bottom case describes interior solutions as a modified markup rule with an added component driven by the bias.

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

The lower expression in Equation 1.3, describes the interior solutions, with the lowest price being P_q . It shows that the optimal price is a modified Lerner equation with an extra term $\left(-\frac{\theta}{1-\theta} \cdot \frac{q_1+1}{1+\epsilon}\right)$. In the interior solutions region, the price is slightly higher than it would have been absent the bias (recall that $\epsilon < -1$). This is due to the left-digit bias making demand less elastic for changes within the same first digit. This added markup is increasing in the bias (θ), and the nominal price (q_1), and decreasing in the absolute elasticity (ϵ).

The pricing schedule is illustrated in Figure 1.2 showing price as a function of cost with left-digit bias. The diagonal gray line is the price without bias, and the thick lines are the optimal prices with the bias, exhibiting regions of costs priced at 99-ending prices and ranges of missing prices with low price endings.

Notice that in Figure 1.2 products whose price is 99-ending are usually (depending on the cost distribution) priced with *lower* markups. This can be seen as a “level effect” micro-foundation of the “image effect” style of explanations – 99-ending prices are indeed, in expectation, of higher value (insofar as costs are positively related to value).

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

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

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

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

The first part of Proposition 1.2 shows that more elastic demand leads to lower next-lowest price. Higher elasticity means lower markups, so costs that go to a 99-ending price q are now higher. If the indifference cost $\bar{c}_q$ is kept fixed, and since $\bar{c}_q = P \left(1 + \frac{1}{\epsilon}\right) + \dots$, it

means that P must decrease. Of course, the cost is increasing, but the proof shows that it increases by less than offsetting this primary effect.

The second part of Proposition 1.2 states that higher bias leads to higher next-lowest prices. The effect of higher bias is simpler to understand. Consider P and increase θ . The infra-marginal benefits from changing the price to the lower 99 are now larger (because the gap increased), while the costs of increasing the price above P are lower because of the lower sensitivity of within-left-digit demand. So P must increase.

In the data we observe prices across time, products and stores. So far, we have not made any assumptions about what drives variation in prices – price variation can be generated from differences in costs, in elasticities, or in bias levels. Like common in the literature (e.g. Hausman (1996)) I will assume that for a specific level of observations (e.g., same brand and chain, across stores and time) elasticity and bias are fixed, and the price-setting variation is driven by changes in costs¹⁴. This assumption is probably more likely for regular prices, and also the appeal of left-digit bias as I will discuss below. Note, that apart from having a support that includes $\underline{c}_q, \bar{c}_q$ (i.e., the costs that generate both a 99-ending price and the next-lowest price above it), we do not need to make any assumptions about the cost distribution, $c \sim F$, to test the NLP predictions in Proposition 1.2.

A simple corollary is showing that the right-digits component of the next-lowest price is increasing from one threshold to the next: For $\epsilon < -1$, the lower digits of the Next-Lowest Price (i.e., $P - P_1$) are increasing with the first digits. This corollary means that more nominally expensive products have larger regions of missing prices. For example, for the same parameters of left-digit bias and elasticity, the lowest observed price beginning with 4 should be \$4.30 and with 5 is \$5.32.

This is not necessarily tied to larger shares of 99-ending prices, because excess mass depends on the shape of the underlying cost distribution. The share of observations in a 99-ending price q is $F(\bar{c}_q) - F(\underline{c}_q)$. Higher left-digit bias or less elastic demand lead to greater $\bar{c}_q$, but both also shift the costs that go to 99, meaning also higher $\underline{c}_q$. It can be shown, even for the simplest uniform cost distribution that the sign of effects on share of 99-prices depends on the parameters. For example, if costs are uniformly distributed, then a change in the excess-mass due to an infinitesimal increase in θ is proportional to $\frac{\partial P}{\partial \theta} + \frac{1}{\epsilon} \left(\frac{\partial P}{\partial \theta} + \frac{1}{(1-\theta)^2} \right)$. If $\epsilon \rightarrow -1$ then this is negative, implying a decreasing mass at q . If $\epsilon \ll -1$ then the first term, which is positive as per Proposition 1.2, dominates, making the effect positive overall. The same kind of exercise holds also for the comparative static effects of the price elasticity¹⁵.

In other words, unlike the next-lowest price predictions, the excess mass is sensitive to the shape of the cost distribution and the exact values of parameters. This means that point

¹⁴DellaVigna and Gentzkow (2017) show that there are different elasticities between stores within a chain-product, but that these differences are not translated to variation in prices; bias heterogeneity is also of course probable, but it seems natural to establish the mean bias as a first step (compare the mean-bias Chetty et al. (2009) to the heterogeneous-bias Taubinsky and Rees-Jones (2016)).

¹⁵For elasticity with uniform cost distribution the first derivative of excess mass at q to changes in ϵ is $\frac{\partial P}{\partial \epsilon} \left(1 + \frac{1}{\epsilon} \right) - \left(P - q + \frac{\theta}{1-\theta} \right) \frac{1}{\epsilon^2}$.

identification of left-digit bias and price elasticity from next-lowest prices and excess mass is theoretically possible, in the sense that the parameters have differing effects on the moments. This avenue is then explored and developed in Section 1.5.

1.3 Data

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

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

Retailer

The first dataset is scanner data of sales of a US national grocer and its subsidiaries, with a relatively rich set of variables. The same data were used in other papers such as Eichenbaum, Jaimovich, and Rebelo (2011), Gopinath, Gourinchas, Hsieh, and Li (2011). Data are collected from the cash registries of the chain's stores and aggregated at the weekly level. Each observation is a distinct product (UPC) in a store in a week. As shown in column (1) of Table 1.1, there are about 74000 distinct products, in 250 stores, and 177 weeks (2004 to mid-2007). The data describe sales of more than 3 billion units, and \$7.2 billion worth of sales.

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

The data are used to conduct demand estimation, and therefore we might be worried about price endogeneity. Ideally we should get prices to vary due to cost changes (supply shocks), but not due to changes in demand. I take a common approach to address this issue, with instrumental variables for prices. I am using leave-out-prices (in the spirit of

Hausman (1996), Hausman, Leonard, and Zona (1994), Nevo (2001)) for the same product at the same period in other stores. To increase the likelihood that these other prices capture supply shocks rather than local demand shocks, I use stores in other cities, that share the same distribution center. This is another advantage of the data, where each of the 250 stores is assigned to one of 17 distribution centers supplying it the goods. I discuss the identifying assumptions with this approach when I describe the estimation in Section 1.4.

An important issue with such scanner data is that if a product did not sell in a certain week, its price is not observed (because zero units sold for zero revenues). It is likely that observations are not missing at random. That is, if a product is less likely to sell when its price is high, then there is a potential for underestimation of price elasticity. To mitigate this issue, I select popular products. I select all products that were available at regular prices for at least 50% of all observations (among weeks and stores in which they were at any point sold) and with at least 10,000 observations. To get some level of minimal price-ending and dollar variation, I am selecting the 1710 products whose regular prices end with 99 at most 75% of observations, or are all within a single dollar digit at most 75% of observations, as shown in column (2) of Table 1.1.

As shown in column 2 of Table 1.1, this procedure leaves me with 1,710 products and 51.3 million observations. Since these are of the most available and popular products, while they are 2.3% of all possible products, they capture 19.2% of all revenue in the data. The average price is \$3.42 (with standard deviation of 1.63), and the inter-quartile range of prices is \$2.29 to \$4.19. Meaning, most of the prices have exactly one dollar digit, limiting exploration of the effects of more or fewer digits (96.35% of observations are strictly between \$1 and \$10).

A potential concern is the sample is selected based on outcome-related variables. Hence, I also present results for the complementary set of popular products, namely those with highly concentrated prices, as a robustness¹⁶. Another limitation of the data is that it is not possible to infer the unit price for products paid by weight, such as fresh produce, meat, or deli items. I exclude these product categories¹⁷.

In order to investigate the relationship of demand to changes within, versus across, dollar digits, the exact price paid is crucial. However, the inferred price (net revenue divided by units sold) does not necessarily, and often do not, describe the actual price paid in store. This happens when some consumers pay with coupons, if some have club membership specific discounts, or if some products are priced with non-linear pricing (such as “buy 1 item get the second for 50%”). Another potential source of multiple prices, which is when prices change

¹⁶The other 2149 products, are characterized by having just as high availability and 10000 observations, but have either more than 75% of observations ending with 99 (only 125 products), or more than 75% of observations starting with the same dollar digit. This mostly happens for cheaper products, priced under \$1. There are 55M observations, representing 170M units sold annually (more than the main sample) with annual revenue of \$292M (less than the main sample). Indeed, the average price is \$2.2 (versus \$3.5 for the main sample), with fewer 99 and 9-ending prices (15.4% and 63.7%).

¹⁷I exclude the following: Delicatessen, Food Service, Fresh Produce, Meat, Seafood, and Alcoholic Beverages. This is a serious issue since weight adjustments only work for some products but not for others. For example, sliced turkey breast, packaged in the deli and priced by weight with a sticker on the wrap will appear as a single unit.

during the measurement week, does not occur in this data set, since the measurement weeks are aligned with the price changing frequency. However, in AC Nielsen RMS data this is a sometimes severe issue as will be described next. The nature of averaged prices in scanner data is a known issue in the literature, and some of its implications on inference regarding economic phenomena, such as demand estimation and price changes, were explored (e.g. Eichenbaum, Jaimovich, Rebelo, and Smith (2014), Einav, Leibtag, and Nevo (2010)). The retailer data have a unique advantage since there is no rounding. e.g., a price in the data, which is the result of dividing revenue by units, might be \$3.044. In such cases, when the price is not “to the cent” then it must be the result of weighted averaging of several prices (here, for example one possible mix can be 41 units selling for 2.99 and 9 units for 3.29). Therefore, I exclude these observations, reducing the number of observations by about 10%, to 46 Million (column (3) of Table 1.1). However, this might keep other observations that are a mixture but happen to have an average that is to-the cent. For example, 9 units purchased at \$4.99 and 6 at \$4.49 will give an average price of \$4.79¹⁸. A different approach is to use the length of the price spell. If the same price appears for two consecutive weeks or more, it is less likely that this price is non-real. Indeed, as shown in Figure A.3 by the dark circles, already at two-week spells, most prices are to-the-cent. This method is mostly useful for Nielsen data, where all prices are rounded to the cent, so the former approach is unavailable. Finally, a sure way to keep only non-averaged prices is to only consider the instances where the net price equals the regular price (which happens for about two thirds of observations). Regular prices do not change very often, and are indeed always to-the-cent.

Some statistics about the price endings are shown in Panel B of Table 1.1. When I include prices that are not to-the-cent I round the prices, emulating what happens in the Nielsen data. As we move from all observations to the “real prices”, the share of 9-ending prices increases, meaning that other price endings are more likely to be the result of averaging of multiple prices. Notice also the higher share of 0-ending prices for all prices (column (3)) versus only regular prices (column (4)). Meaning, 0-ending prices are strongly positively correlated with being on-sale¹⁹.

If 9-ending prices are mostly not on-sale, while 0-ending prices (and others) are on-sale, it will lead to spurious higher demand for 0-ending prices without flexible enough controls of sales effects or if the model is misspecified. Prices are only one determinant of demand, and different promotion techniques can lead to stronger effects than others. Indeed, as casual observations suggest, round prices are commonly the result of quantity sales (e.g., “2 for \$5”

¹⁸Under the assumption that there are at most two prices paid, with one of the prices being the regular price, the other price and quantity purchased at each price can be inferred in theory from the average price, the total number of units, and the facts that prices should end with a cent. However, this is imperfect and allows for multiple prices, and requires some assumptions to restrict possible prices to a single price. I therefore prefer to err on the safe side by reducing sample size.

¹⁹Note that these findings stand in contrast with the assumptions of “image effects”. First, under image-effects rationalization, round prices should be linked to high-quality items, but at the product level, they occur mostly when the product is on-sale. Second, a common explanation for 99 prices effectiveness in this literature is that they are learned through consumer experience to signal a discount. But, the contrary is true, as 9-ending prices, if anything, should signal the regular, higher, prices.

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

Nielsen

Nielsen Retail Scanner (RMS) and Consumer Panel (HMS) data are provided by the Kilts Center at the University of Chicago. Like the retailer data, RMS records weekly UPC-store level quantity and revenues. Overall the scanner data record observations for over 35,000 stores during 2006-2014 in the US, regarding about a million different unique products. As per this version, I am using the Nielsen data as a robustness test using a sample of products and stores selected by DellaVigna and Gentzkow (2017) (see for details on data selection), of 12 items (termed “high-end” in their paper) in food stores. The summary statistics are described in Appendix Table 1.1. The sample consists of 28 million observations of these 12 products, in 11000 stores, with a total of 188 million units sold for \$881 million. The data have a few disadvantages relative to the retailer data, which I elaborate on with possible solutions in Appendix Section A.4. In a nutshell, an effective way to identify sales and non-real prices in the data is by using the spell length of a price. If a popular product’s price changes from week to week, it is likely to be a non-real price. If a product’s price is kept constant for many weeks, e.g. for six consecutive weeks, it is likely a regular price. I discuss this idea further in Appendix A.4.

1.4 Demand Estimation

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

Reduced-Form Evidence

Figure 1.3(a) illustrates constant elasticity demand curves with and without the bias, i.e. $\log Q = A + \epsilon \log((1 - \theta)p + \theta \lfloor p \rfloor)$, for $\theta = 0.15$ on the left panels and $\theta = 0$ on the right. On the right, without bias, demand is a line with slope ϵ . On the left, the slope is flatter within each dollar digit, and demand drops when the dollar digit changes (highlighted by different

colors). The dashed line is the fitted line from a $\log(\text{quantity})$ on $\log(\text{price})$ regression²⁰. In turn, Figure 1.3(b) shows the residuals by price from these regressions. The sawtooth pattern of residuals for the bias case, with discrete drops at dollar thresholds, is a manifestation of the bias. If there is no bias, as in the right panel of figure 1.3(b), residuals are a flat line at 0. If the model is misspecified, it will exhibit some other general shape.

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

Figure 1.4 shows that the sawtooth patterns are found in the data. Residualized demand is increasing within a dollar-digit and drops across digits. But before turning to the results, I would like to make the empirical strategy clear.

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

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

The third challenge is the sorting of *price-endings*. As long as prices change due to costs, and as long as there is enough variation in price-endings (guaranteed by sample selection) this is not a worry. The reason is, that I am estimating the demand curve, and while prices

²⁰The simulated price distribution is uniform. If it were the case that prices are more likely to be 99 ending, a regression line will be closer to the 99 ending prices. The slope of the regression line, without taking the bias into account, is interpreted as the price elasticity. The slope is affected by the distribution of price endings, but can go either way. For example, if prices are \$4.99 and \$5.05, the regression slope will be steeper, making demand seem more elastic than it is, while if prices were \$5.05 and \$5.99, the slope will be flatter and demand less elastic. I assume, as in the model, that the price elasticity is the demand response to true price changes.

are a function of the costs, demand is not. From the consumer perspective, it does not matter that different costs translate to the same price. Meaning, the shape of the demand, by price, is not affected by sorting on price. Of course, if all prices end with 99 or all prices begin with the same dollar digit, there is no variation and no hope for separating the price-ending effects from price effects.

Since the data lack exogenous variation in prices, we worry of elasticity bias driven by unobservables (such as local demand shocks) affecting prices and quantity sold. However, much of the demand variation is absorbed by fixed effects, capturing seasonality, products characteristics, permanent store differences, and sale strategies. As Rossi (2014) argues, long panels with multiple fixed effects might not be subject to substantial bias²¹. Relatedly, DellaVigna and Gentzkow (2017) show that chains set uniform prices across stores, hence implying that price variation is not driven by local shocks and does not equate demand and supply at the store-week level. Meaning, the worry should not be too big. So, even though I use instrumental variables for the main results, in some cases I estimate a simple OLS (or a non-instrumented NLLS). I construct instrumental variables and find slightly more elastic demand, but small impact, if any, on the patterns and estimated bias. I instrument for $\log P_{ist}$ with $\log P_{id(s)t}$, the average log-price of the same item i in same period t in stores (s) in other cities but the same distribution center $d(s)$. The assumption here is that this average price is capturing changes in price that are driven by cost changes. The costs are likely to be common across stores and hence are unlikely to be driven by local demand shocks²². These instruments can not fully resolve the issue of endogeneity, since it might be that there are aggregate demand shocks that affect prices across stores in a certain week.

Finally, cross-elasticities with other products are an important component affecting demand²³. I am therefore including a variable of the average price of other products in the same category (as defined by the retailer, e.g., “tea”). The common approach in the literature is to add all “close” products flexibly (as in the lowest level of almost ideal demand systems, as in Hausman et al. (1994), or to use BLP-like random coefficients (as in Nevo (2001)), or to take a Bayesian approach (Rossi, Allenby, and McCulloch (2012)). Since I am using thousands of products and tens of millions of observations, and since I am excluding some observations with non-real prices, these approaches are harder to implement in practice. By adding a single variable of mean log price, I essentially estimate the average cross-elasticity to products within the category (I find it to be on the order of 0.1-0.3). I also have a robustness

²¹In fact, Hausman (1996) argues the following: “to the extent that supermarkets set their prices... under an assumption of constant marginal costs (in the short run) and do not alter their prices to equilibrate supply and demand in a given week, prices may be considered predetermined... then IV methods would not necessarily be needed”.

²²The first stage for all products is very strong, with the mean coefficient across products being 0.86 (standard deviation of 0.14) with the smallest F-stat being 69 and most are on order of tens thousands. This strong relation is likely driven by zone pricing behavior, and also explains why a simple OLS provides almost identical results, just less noisy.

²³Cross elasticities are also important for pricing. To keep the model solvable analytically, and allow for comparative statics, I ignored cross-elasticities on the theory and estimation of the firm behavior. I add it back in conducting counterfactual exercises.

specification with a category price decile fixed effect to get non-linear effects.

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

$$\log Q_{ist} = \epsilon_i \log P_{ist} + \alpha_i \log P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu_{i,on-sale(ist)} + \mu'_{i,spell.length(ist)} + \nu_{i,cents.digit(ist)} + e_{ist} \quad (1.6)$$

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

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

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

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

The main results, for 1,710 products are shown in Figure 1.4. The horizontal axis is the price, and the vertical axis is the residualized demand. Each point is the estimated $\alpha_{\tilde{p}}$, the residual demand at a specific price (in 10 cent bins). The top panel shows the residuals from specification 1.6 on the main sample (column 3 in Table 1.1). The bottom panel shows results from a restricted sample of regular prices, and the patterns are cleaner there (column 4 in Table 1.1). The figures show that quantity purchased, netted of price, promotion, and seasonality effects, is increasing within each dollar-digit and drops at dollar crossings. Such patterns are consistent with demand of left-digit biased consumers. Meaning, after controlling for all other effects, higher price-endings are associated with higher demand. Under the model, this is driven by ignoring a fixed proportion of the price-ending, creating larger discounts of the total price for higher price-endings.

In the main sample, for some prices there is actually no drop crossing the dollar threshold but for the next 10-cent prices (for example, demand only drops at \$5.10 rather than at \$5.00). A simple comparison of demand at \$4.99 to \$5.00 would lead to the conclusion that there is opposite bias, suggesting that perhaps round prices have some innate appeal.

However, as can be seen in the bottom panel focusing on the sample of regular prices, these patterns are driven solely by promotions. For the restricted sample of regular prices, the round prices are aligned with the overall sawtooth pattern and all drops are at round prices. It does mean, however, that despite my efforts, controls for on-sale effects on demand are not fully absorbed. Or, more fundamentally, that the model is more suited to describe price effects absent promotions.

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

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

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

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

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

In appendix figures A.5, I explore the role of the demand structure. The patterns are not driven by the constant elasticity assumption, since the figure shows that a semi-elasticity specification produces similar sawtooth patterns. To estimate the patterns, I use the same specification and procedure as in Equation 1.6, replacing log-price with actual price (also replacing average log-price of other products with average price, and instrumenting with average price in other stores). In Appendix Figure A.6 I do the same exercise but instead regress log-price on a 5th degree polynomial in prices at the product level, which is a very flexible, yet smooth, demand structure. In panel (b) I control for the possible threat of the retailer choosing 99-ending prices when they want to promote a certain product in a category. I add controls for the within-category weekly price decile. For example, being in the cheapest 10% of dairy products is absorbed by the fixed effects.

Appendix Figure A.7 provides robustness tests on a different sample, to show that these

patterns are not driven by product selection. The figure shows the residuals from instances with regular prices for all remaining 2035 popular products (defined by at least 50% availability at regular prices and at least 10000 observations), with at least 75% of prices ending with 99 (merely 125 products) or at least 75% of prices starting with the same dollar digits. Meaning, popular products, but those with more concentrated price distribution, still produce the same sawtooth patterns.

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

Structural Estimation of θ

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

As a benchmark, I estimate a homogeneous left-digit bias parameter with a semi-elasticity specification (details in section 1.4). I find a precise estimate of $\theta = 0.1635$ (0.0096) which means that on average there is a 16 cent perceived difference between a 99-ending price and the 1-cent higher round price²⁴.

For the main analysis, similar to the reduced-form results, I estimate the left-digit bias parameter at the product level. That is, I estimate multiple θ_i s. I do so both for computational reasons²⁵, and in order to explore heterogeneity in the bias estimates. Some of the heterogeneity is coming from sampling bias. Even with tens of thousands of observations per product, there is only a small finite number of prices used, while for identification, multiple prices within and across dollar digits are required. Further, the specification is a simplification of the demand structure, especially lacking in controlling for promotion effects. But, this heterogeneity is strongly correlated with underlying characteristics of the different products and their clientele. Learning about this heterogeneity structure informs us about the underlying mechanisms driving the bias in the field.

²⁴The specification is $\log Q_{ist} = \eta(1 - \theta)P_{ist} + \eta\theta \lfloor P_{ist} \rfloor + \alpha \log P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \mu'_{i,on.sale(ist)} + \nu_{i,cents.digit(ist)} + e_{ist}$. Errors are clustered at the store level. Standard errors on θ are calculated using the delta method.

²⁵The above single regression takes almost two weeks to run because of the about 470,000 fixed effects. For the NLLS method, it is impractical for a numerical algorithm to be able to handle so many parameters.

Sufficient Statistics Approach

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

In general, for each triplet of demand drop $\Delta \log Q$, elasticity ϵ , and dollar threshold d , the bias is²⁶:

$$\theta = d \left(1 - e^{\Delta \log Q / \epsilon} \right) \approx -d \cdot \Delta \log Q / \epsilon \quad (1.7)$$

The first estimation approach relies on this specification, where I estimate the demand drops for each product-dollar pair, and complement it with the elasticity from the IV specification in Equation 1.6. To get the demand drops I estimate for each product

$$\begin{aligned} \log Q_{ist} = & \sum_{\{d=\text{dollar-digit}\}} (\epsilon_{id} \log P_{idst} + \kappa_{id}) + \\ & \alpha_i \log P_{c(i)st} + \beta_{is} + \gamma_{i,\text{year}(t)} + \delta_{i,\text{month}(t)} + \\ & \mu_{i,\text{on-sale}(ist)} + \mu'_{i,\text{spell.length}(ist)} + \nu_{i,\text{last-digit}(ist)} + e_{ist} \end{aligned}$$

clustering standard errors by store. Then I calculate 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. As shown in Panel A of Table 1.2, I find an average drop of 9.0 percentage points (median of 8.0), where 86.4% of drops (keeping the most precisely estimated drop per product, in case there is more than one) are positive, and 54.5% of drops are significantly positive at the 5% level (compare with 2.8% significantly negative). The drops are a product-level econometric equivalent of the residuals figure, showing that this phenomenon is occurring product by product and not only at the aggregate.

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

²⁶This is an approximation of $\Delta \log Q = \epsilon (\log((1 - \theta)d + \theta(d - 1)) - \log(d)) = \epsilon \log(1 - \frac{\theta}{d}) \approx -\epsilon \frac{\theta}{d}$ for $\theta/d \ll 1$.

The drops, then, provide a reduced-form measure of the effect of a left-digit change. This is similar to Anderson and Simester (2003) or Macé (2012), who add a dummy for any 9-ending price. This exercise adds to that existing knowledge by doing so for many products over a large set of stores and chains. Further, combining with elasticity estimates gives a structurally meaningful measure – the left-digit bias is the ignored share of the last digits. This estimation approach, relying on Equation 1.7, has the advantage that the vertical drops are unbiased estimates even if there is price-ending sorting due to the bias, as long as there is no systematic sorting within product to price-endings by demand shocks. The elasticity estimates are unbiased to the extent that controls and instrumental variables solve endogeneity issues. Therefore, the left-digit bias estimates are arguably unbiased.

A histogram of estimates is shown in Figure 1.5. Table 1.2 shows that the mean parameter is 0.26, with inter-quartile range of 0.07 to 0.39. For the figure, I am highlighting in darker shade the estimates where the estimated drops are in the top three quartiles of precision. As can be seen in the figure, imprecise estimates generate most of the outliers. A CDF of all estimates is shown as the solid dark curve in Appendix Figure A.8.

Non-linear Least Squares

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

$$\begin{aligned} \log Q_{ist} = & \epsilon_i \log((1 - \theta) P_{ist} + \theta \lfloor P_{ist} \rfloor) + \\ & \alpha_i \log P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \\ & \mu_{i,on.sale(ist)} + \mu'_{i,spell(ist)} + \nu_{i,cents.digit(ist)} + e_{ist} \end{aligned} \quad (1.8)$$

To estimate it I code the analytical gradient and hessian, and minimize the sum of squared errors with a numerical optimizer. I get the initial guess from running an OLS without the bias, and guessing a θ of 0.2 (I get the same results for different initial values of θ). The identifying assumption being that variation in prices and in price-endings is not driven by changes in elasticity. I am relying on variation in prices within and between dollar digits, absorbing the variation across years, month of year, stores, sale effects, and last digits with fixed effects. In the NLLS I do not instrument with other prices, and to get standard errors I use cluster bootstrap, sampling different stores with replacement. The CDF of product level estimates is the dashed blue curve in Appendix Figure A.8, which is very similar to the estimates coming from the drops. As shown in Table 1.2, the mean bias is 0.24 and the median is 0.18.

Semi-elasticity

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

$$\begin{aligned} \log Q_{ist} = & \eta_i (1 - \theta_i) P_{ist} + \eta_i \theta_i \lfloor P_{ist} \rfloor + \\ & \alpha_i P_{c(i)st} + \beta_{is} + \gamma_{i,year(t)} + \delta_{i,month(t)} + \\ & \mu_{i,on.sale(ist)} + \mu'_{i,spell-length(ist)} + \nu_{i,cents.digit(ist)} + e_{ist} \end{aligned} \quad (1.9)$$

where η is the semi-elasticity (the change in log quantity due to a unit change in prices), and the ratio of estimated coefficients on the price and the floor of the price gives the ratio of $\frac{1-\theta}{\theta}$. Reassuringly, the estimates from that approach are similar to the estimates from the NLLS approach, as can be seen in Figure A.8. The estimates have a correlation of 0.85 and a rank correlation of 0.91. As shown in Table 1.2, the mean bias is 0.20 with a median of 0.16.

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

Using a similar specification, I estimate the model with homogeneous semi-elasticity and left-digit bias across products (but with product level fixed effects and cross elasticities)²⁷. I estimate a left-digit bias for these 1710 products of 0.1635 (0.096) and semi-elasticity of -0.3697 (0.0044). Since the mean price is \$3.50 it implies an average elasticity of -1.29.

Exploring Bias Heterogeneity

The demand-side model makes no assumption about where the bias is coming from and what mental process it might represent. One immediate question to ask is what is correlated with the bias (see also Macé (2012)). The results provide field evidence of relative thinking that co-lives with a form of deliberate inattention. Because this is not at the core of the paper, I defer the analysis and discussion to Appendix A.5. In a nutshell, utilizing the Nielsen Consumer Panel (HMS) to gain characteristics of each product and the people who consume it, I find that the bias is increasing with the price, decreasing with the number of units a household purchases in a shopping trip, and is not significantly correlated with the purchasing clientele's education or income.

²⁷Specifically, I estimate

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

1.5 Firm Pricing

Building on the findings of the demand structure in Section 1.4, I turn to investigate the firm response. I document that qualitatively, firms act as if they respond to left-digit bias. And since they seem to respond to the bias, I conduct an “as if” exercise, estimating the perceived left-digit bias from their pricing behavior. I find large gaps between the demand-side estimated bias and the supply-side estimated perceived bias.

Qualitative Predictions

As analyzed in section 1.2, the model prescribes three key qualitative predictions for pricing patterns of firms facing a demand structure such as the one estimated in section 1.4. The model predicts for optimizing firms: (1) excess mass at 99, (2) which is coming from a region of missing prices with low price-endings, (3) and increasing with the dollar digit.

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

Excess mass Figure 1.6 shows the price-ending histograms for the regular prices in the retailer data and the long price spells in the Nielsen data (at least 6 weeks). The reason I restrict attention to regular, or likely regular, prices, is that as shown above promotions techniques actually might work to increase demand better with round prices (see also Anderson and Simester (2003)). This is possibly because such pricing startegites, such as quantity discounts (“2 for \$6”), tap on different psychological forces such as price fluency and suggestive quantities. These forces are beyond the model. Indeed, the model seems to fit best for the regular prices as in Figure 1.4(b).

High shares of prices indeed end with 99, 27% in the retailer data, and 34% in the Nielsen data. Also, almost all prices end with 9 (85% and 86% percent)²⁸. This latter finding also implies that there is further bias ignoring the last digit of the price, hence leading to so many 9s. This effect is probably true, but impossible to estimate, exactly because there is lack of variation in the last digits, and hence the drops in demand at the dime thresholds can not be identified. However, I do robustness tests with additional bias regarding the last digits (that is, assume that with some measure θ_2 consumers see the price as the dollar and dime digits but ignore the cents).

²⁸In 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%).

Missing prices Figure 1.6 also shows that for both the retailer and the Nielsen data, there are few prices ending with anything lower than 19. This finding has not been described, predicted, or explored in prior papers. Meaning, this finding provides further support for the model, especially because it is an unanticipated pattern and prediction.

It also does not seem that retailers price at 99 out of “convenience” or as if they perceive there is an inherent higher demand at 99 versus any other price (e.g., a “blip” in demand as discussed above). Such behaviors are manifested in symmetric missing prices, which is not the case here^{29,30}.

Next-lowest price by first digit Figure 1.6 masks the heterogeneity of how price endings differ across dollar digits. The model predicts that the next-lowest price should increase with the dollar digit, keeping the bias and elasticity constant. If, as we have seen above, the bias is increasing with the price, this prediction intensifies. The share of 99-ending prices is not clearly predicted, since it relies on the distribution of costs.

In order to examine how price-endings change across dollar digits, I display the left tail of a by-dollar CDF of price endings in Figure 1.7. Darker lines are higher dollar digits. Indeed, Figure 1.7 shows that the distribution of price-endings is shifting to the right, meaning that for higher dollar digits there are less low price endings.

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

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 1.4?

Recall from Proposition 1.2 that next-lowest prices are decreasing as demand is more elastic and as the bias is weaker. Therefore, I take the low-end estimated bias level, of 0.16, and the mean elasticity from the IV specification of -1.5³¹. For these levels, the predictions are that next-lowest prices are 99. In that case there is no ambiguity, *all* prices should end at 99 regardless of the cost distribution.

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

²⁹In a related paper, Dube et al. (2017) study wage offer distributions. If employees are left-digit biased and want higher wages, then wages should be set at round numbers. The authors find excess mass at 0-ending wages, but missing prices are symmetric. They use the symmetry to argue against left-digit bias in their setting

³⁰One possible explanation for the gap, is that they used wages in cents in their experimental design, below and above 10 cents, but Hilger (2018) finds no bias at these levels, while finding it at the dollar, tens and hundreds of dollars digits.

³¹Since I am comparing the *regular* prices distribution to the estimates using all prices, it is worth noting that if I estimate the bias parameters using observations of regular prices only (with NLLS), I get slightly higher estimates, with median bias of 0.23.

assume that the retailer is pricing as described by the model, according to what they *perceive* to be the price elasticity and left-digit bias. In other words, in Section 1.4 I estimated the elasticity and bias that rationalize demand behavior, and here I estimate the elasticity and bias that rationalize supply behavior.

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

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

The idea behind the identification is that left-digit bias creates discontinuities, thus making frictions in pricing that can be exploited for identification. In a sense, this is the dual of the notch estimation in the taxation literature (see Kleven and Waseem (2013), Saez (2010))³². 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 estimands³³. In addition, the price distribution is more discrete than just masses at 99, as almost all prices end with 9. If this is driven by added left-digit bias to the cent digit, different parameters would translate to the same 9-ending next-lowest price, tampering with

³²In Public Economics the idea is that non-linear tax schedule creates incentives for tax payers to change their reported income (or labor supply) and bunch at kinks or notches of the tax schedule. The duality can be seen as follows: budget set of taxes is like the demand (elasticity); the notch given by the tax schedule is like the bias parameter; the utility function is like a profit function; and the latent ability distribution is like the cost distribution - estimated from the observed income distribution like the observed prices outside of the bunching and missing prices.

³³Kleven and Waseem (2013) face a similar issue where no one should have income in dominate regions. They interpret it as frictions, and implicitly assume that frictions are random. The approach I follow is implicitly taking a similar stance, since I put equal weight on many price moments.

the variation needed for identification.

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

To satisfy the need for a smooth cost distribution, I am aggregating over products, assuming a constant left-digit bias and elasticity. Which leads me to the next caveat - the assumption of homogeneous parameters. Heterogeneity of the bias parameter can of course lead to low-ending prices, and is definitely plausible, as shown in Section 1.4. However, identification with heterogeneous bias relies far more heavily on the shape of the cost distribution, which is an arduous requirement to infer from the data. I argue that since there is such a stark difference in predictions on aggregate and product-by-product (there are only 11 out of 3859 popular products whose prices end solely by 99), assuming a single left-digit bias parameter is a reasonable first step. In addition, in the next section I explore the implications of heterogeneous parameters on pricing and profits, and find them to be small.

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

Let S_p be the share of prices set at a price p . I use minimum-distance estimation to match the empirical price densities, $\{S_p\}$, to the predicted price densities, $\{\hat{S}_p\}$ (where p belongs to a grand price vector $\mathcal{P}$). For example, take $\mathcal{P} = \{2.29, 2.39, \dots, 3.69\}$. The algorithm is as follows: (1) Fit a polynomial to the empirical price distribution using prices that are not bunching at 99 or missing due to the bias, call that price distribution $\hat{F}_p$. Then, for each pair of elasticity ϵ and left-digit bias θ : (2) take all relevant 99-ending prices, $\{q_i\}$ and calculate the matching next-lowest prices $\{P_{q_i}\}$ using Equation 1.5 (for example, say $\{q_i\} = \{1.99, 2.99\}$ and that $\{P_{q_i}\} = \{2.51, 3.53\}$). (3) expand the vector of prices to include $\{q_i\}$, $\{P_{q_i}\}$, and all 9-ending prices between each pair of P_{q_i} and q_{i+1} , giving a vector of prices $\mathcal{P}'$ (e.g. $\mathcal{P}' = \{1.99, 2.51, 2.59, 2.69, \dots, 2.99, 3.53, 3.59, 3.69\}$). (4) Given θ and ϵ calculate for each $p_i \in \mathcal{P}'$ the cost c_{p_i} , such that p_i is the profit-maximizing price. By construction this is given by the first order condition as $c_{p_i} = p_i \frac{1+\epsilon}{\epsilon} + \frac{\theta}{1-\theta} \frac{\lfloor p_i \rfloor}{\epsilon}$. (5) Then for each price, the share of observations at that price are $\hat{S}_{p_i} = \hat{F}_c(c_{p_{i+1}}) - \hat{F}_c(c_{p_i})$. And we get $\hat{F}_c$ using the following model-induced identity:

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

(6) expand $\mathcal{P}'$ to $\mathcal{P}$ by adding the missing prices, between q_i and P_{q_i} , and assign them predicted zero shares³⁴.

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

This procedure might be somewhat sensitive to which moments are chosen, what moments are excluded from the fit estimation, and the degree of the polynomial fit. I therefore take 32 combinations of estimation settings³⁵. For example, in one of these settings moments are between \$1.29 and \$5.89, using all prices with price endings between 29 and 89 for the $\hat{F}_p$ fit, with a polynomial of degree 3. I minimize the distance of actual moments from predicted with price endings strictly higher than 89 and lower than 29 (that is 99, 09, and 19), since these are the prices that were not used to predict the price distribution³⁶. Then, for each setting I run 300 cluster-bootstraps at the product level.

Table 1.3 presents the estimation results showing the estimated perceived left-digit bias and price elasticity for the main sample, using regular prices of the 1710 products. I am guessing 50 initial values of bias and elasticity and choose the estimation with the best fit. To get confidence intervals I run additional thousand cluster-bootstraps at the product level (each with 50 initial guesses). I am estimating homogeneous parameters and find a mean perceived bias of $\hat{\theta} = 0.023$ and elasticity of $\hat{\epsilon} = -3.35$. I can reject a perceived left-digit bias larger than 0.043, but can not reject that the price elasticity is -1.5.

If I restrict the elasticity to be -1.5 and only estimate the bias I get tighter estimates (with similar goodness-of-fit) for the same bootstrapped sample of $\hat{\theta} = 0.0047$. The similar goodness-of-fit between the unrestricted and restricted versions means, as predicted by the comparative statics results, that there is weak identification separating between the elasticity and left-digit bias parameters. However, the firm is underestimating the bias at least by a factor of 3, and more likely by an order of magnitude. As a robustness, a break-down of the estimates by specific setting with 95% confidence intervals per each setting is plotted in Figure 1.8. The figure shows that while point estimates are sensitive to the moments choice and the assumed shape of the cost distribution (via the polynomial fit), confidence intervals are stable.

³⁴relabel non-9-ending next lowest prices as the lower 9-ending price, without changing the share. For example, 2.51 will be relabeled as 2.49 with a share $\hat{S}_{2.49} = \hat{F}_c(\mathcal{C}_{2.59}) - \hat{F}_c(\mathcal{C}_{2.51})$

³⁵I take all combinations of

$$\begin{aligned} & \{\text{bottom} - \text{cents}\} \times \{\text{top} - \text{cents}\} \\ & \times \{\text{lowest} - \text{dollar}\} \times \{\text{highest} - \text{dollar}\} \times \{\text{polynomial} - \text{degree}\} = \\ & \{0.29, 0.59\} \times \{0.79, 0.89\} \times \{\$1, \$2\} \times \{\$5, \$6\} \times \{3, 7\} \end{aligned}$$

³⁶Appendix Figure A.12 shows an example of the actual moments, highlighting which were excluded from the fit to create the counterfactual distribution in light yellow. The dark bars are the shares that are being matched by the predicted moments. The predicted moments are shown in semi-transparent bordered bars.

The two rightmost columns of Table 1.3 are showing the results using the alternative sample of popular products, namely those with less variation in prices (out of the 2149 products, 31 have shares of 99-ending prices above 75%, and the rest have most of their prices concentrated within a single dollar digit). The same pattern arises, with very similar estimates for both samples, and similar patterns of a flat goodness-of-fit between the restricted and unrestricted specifications. This is useful since there might be a worry that by choosing the 1710 products that exhibit less variation in prices leads to a mechanical finding of lower perceived bias (even though the demand side estimates should be consistent with their pricing if the bias is perceived correctly). Finding similarly low perceived bias for the products with more 99s and less between dollars variation adds to the robustness of the results.

This method can be used to estimate the price elasticity from a price distribution in other settings, especially if one is willing to take a stance regarding the left-digit bias. Meaning, if the researcher does not have access to quantities sold nor costs but observe multiple prices for similar products with fluctuating costs, they can estimate what the firm perceives to be the price elasticity. Appendix Figure A.13 shows the estimated elasticities for the various specifications.

1.6 Counterfactuals

Left-digit bias and how firms respond to it affect welfare. Left-digit bias of consumers has direct effects on *demand*, and the perceived left-digit bias of firms has effects on *pricing*. Together, these forces affect firm profits, consumer surplus, and deadweight loss. I examine two counterfactuals. First, I assume firms optimize and study the effects of left-digit bias versus a no-bias world. Second, I study underestimation of the bias by the pricing firm. The main findings from these exercises are that, (1) by lowering prices, left-digit bias has ambiguous effects, and might even increase consumer surplus and lead to a more efficient outcome; (2) firms lose a few percents of gross profits by underestimating the bias with small effects on consumer surplus.

The effects of the bias on demand hinge on the discontinuities at round prices, but also directly depend on the level shift in represented by Δ (recall $\hat{p} = (1 - \theta)p + \theta(\lfloor p \rfloor + \Delta)$). As corner examples, when $\Delta = 0$ demand, and hence profits, are higher everywhere, while if $\Delta = 0.99$ demand and profits are lower. Clearly, the first order effects on profits are driven by Δ . However, Δ can not be inferred from the demand nor supply estimations. Therefore, I focus on comparative statics conditional on Δ .

Unlike the demand effects, pricing decisions are a function of what the firm perceives the bias to be, $\hat{\theta}$, rather than the bias itself, θ . A price governed by $\hat{\theta} > 0$ versus $\hat{\theta} = 0$ can be higher (e.g. 4.29 instead of 4.20) or lower (e.g. price will go down from 4.20 to 3.99). These mixed movements of prices are key to understanding the overall effects. Importantly, these are functions of θ and the elasticity, and are not very sensitive to Δ ³⁷.

³⁷The force pushing prices up are the lower elasticity within a dollar-digit price change, while the force pushing prices down is the discontinuity in demand at round prices. Both of these forces exist regardless of

Together, the demand and pricing effects lead to distortions in consumers choice and firms performance. I will now examine the effects of bias existence when firms respond to it optimally, and then the effects of firms underestimating the bias.

The effects of left-digit bias $\theta > 0$

First, consider the case when the firm prices optimally according to the true level of the bias. The equilibrium outcomes are illustrated in Figure 1.9. The key idea in interpreting the figure is that welfare is governed by the true demand curve, prices, and costs. Assume first that $\Delta = 0$ as in panel (a). Compared to the outcome with no bias, point x , the equilibrium price with bias $\theta > 0$ can be the 99-ending price below, which is point y . In that case, consumers enjoy the lower price $p(y)$ and purchase more of it, leading to an increase in consumer surplus represented by the gray trapezoid; but consumers over-consume because $\hat{p}(y) < p(y)$, leading to a transfer of surplus from consumers to firms, represented by the triangle. Therefore the overall effect of a lower price depends on which one dominates. Further, the price might actually go up, as in y' , in which case consumers are clearly worse off. Deadweight loss is also ambiguous in sign since the overall quantity sold goes up for prices at 99-ending prices (as in y) but goes down if prices increase due to the bias (as in y').³⁸

As shown in Section 1.2, the left-digit bias pushes some prices up to y and others down to y' . The average change depends on the exact demand elasticity and underlying cost distribution, but is fundamentally mixed. Even if the average price across products does not change, due to non-linearity of demand there can be aggregate effects on quantity sold, and on deadweight loss and consumer surplus.

In contrast to consumer surplus, one effect is easily determined, which is the effect on producer surplus as mentioned above.

Because Δ is unknown I simulate the effects under various assumptions. Indeed, effects vary in sign and magnitude when assuming different Δ s (see Appendix A.6 and Appendix Table A.4). Further, even fixing Δ , the effects of higher θ can be non-monotonic. Meaning, the possibility of mixed effects is easily found in simulations.

Overall, even though it might be expected that being biased will necessarily harm consumers and lead to less efficient outcomes, that is not the case. Consumers bias creates opportunities (or constraints if $\Delta = 0.99$) for the firm to gain from lower prices, and as such can effectively cut down monopolistic margins. It does come at a cost of consumption distortions, but the net effect, on both consumers and deadweight loss, is inherently undetermined due to the opposing changes in price.

Δ even though their magnitude is affected to some degree.

³⁸The effects in panel (b), when $\Delta = 0.99$ are similar. moving from x to y leads to an increase in CS, here even unambiguously because the 99-ending perceived price is closest to the truth. Still, a change from x to y' leads to a decrease.

The effects of firms underestimating the bias $\hat{\theta} < \theta$

Now consider a scenario, as I find in Section 1.5, in which firms underestimate the bias. The extreme version, in which the firm ignores the bias is captured in Figure 1.9 as point z . The price is the same as in point x , but the quantity shifts (up in panel (a), down in panel (b)) due to the biased demand. That is, if there are no price effects, the distortion only harms consumers as it distorts consumption choices away from the optimum. In contrast, an underestimated bias that is not zero, will mean that the price will be more likely to be on the curve between (z, y') - there will be fewer instances in which the price will go to 99-ending y , and when corrected up will be kept lower than in y' . As in the previous discussion, the effects on consumer surplus and deadweight loss are ambiguous, for we are doing a similar comparison between x and y or y' .

It is clear then that the firm is in a worse situation by pricing at z , which is not one of the possible profit-maximizing options y and y' . But perhaps this is not a costly mistake. For that reason I will now quantify the profits lost by underestimating the bias. Using the estimates from the demand side as “the truth”, I compare profits under different pricing strategies: if the firm is ignoring the bias, fully optimizing, or pricing by the perceived bias as estimated in Section 1.5³⁹.

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

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

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

The results are shown in Table 1.4, where I present the gross profits lost by pricing according to $\hat{\theta}$ versus the profits if the firm were to price according to θ . As a sensitivity test, I assume different distributions of the elasticity, left-digit bias, and perceived left-digit bias parameters. Panel A shows that if the firm were to ignore the bias altogether, that is price as if $\hat{\theta} = 0$. Panel B shows the loss if the perceived bias is as estimated, lower than the actual but not ignored. In Appendix A.6 I conduct a similar exercise, where instead of assuming a constant elasticity demand, I simulate various demand structures stemming from populations with various latent product valuations. I find similar magnitudes of profits forgone as in the constant elasticity exercise (see Appendix Table A.4).

I take the point estimates from the demand and supply side as benchmark, and estimate that the retailer is losing 2.7% of profits versus the optimum, out of a possible loss of 4.3% if were to ignore the bias. If I assume that the bias is actually lower, at 0.1, that the elasticity

³⁹The reason I am focusing on profits lost is because the sign of these effects is not driven by assumptions, while other effects (CS and DWL) are sensitive to assumptions, as shown in Appendix Table A.4.

is on the high end from the perceived estimation, at -4, and the perceived bias is also higher at 0.03, I still find losses of 1.1% of gross profits. I then explore the effect of heterogeneity of each of the parameters, and it does not carry a large impact on the results. Finally, I also estimate a case of a 2-product monopolist with an elasticity of substitution of 0.1 between them. For the particular distribution the potential loss is 3%, and the perceived bias leads to a loss of 1.7%.

In Appendix Table A.3 I add two robustness tests for each of the scenarios. In the first I assume that consumers also ignore the last digit with $\theta_2 = 0.25$. It bears a very small effect on the results. The second robustness is more fundamental. I assume that the bias is such that when consumers ignore the lower digits they replace them with .99 instead of .00. It decreases the losses to be 1%-2% across scenarios⁴⁰.

Across scenarios and assumptions I estimate that the retailer may lose 1.1%-3.2% of gross profits. If, as in Montgomery (1997), operating margins are about 12%, it means that the firm is losing 9%-27% of operating profits by underestimating the bias. Interestingly, following circulation of a previous version of this paper, Hilger (2018) conducted the same empirical exercise on proprietary online subscriptions of private vendors, and finds similar effect sizes.

The losses are coming from pricing at dominated regions. Note, also, that as Appendix Figure A.10 shows, the overall price level actually does not change much, and if anything is somewhat lower. This result is insensitive to Δ as shown in Appendix Table A.3. Namely, it does not matter if demand is overall higher or lower, what matters is that pricing at dominated regions is a costly mistake.

1.7 Conclusions

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

Two main explanations for the discrepancy are that either the model is missing something fundamental, or the firms are making a consistent mistake.

⁴⁰Considering CS effects in the same pricing scenarios I find that consumer surplus varies a lot with different assumptions. The effect of firms underestimating the bias changes consumer surplus from -4.6% to +2.8% compared to if firms were to price according to the true bias. However, pricing with underestimation versus pricing optimally in a world with no bias at all produces consumer surplus level that are almost the same ($\pm 0.1\%$). This is merely suggestive, as under other assumptions the effects can amount to a few percents.

First and foremost, the pricing model is very simple. It simplifies competitive forces to enter through the price elasticity. Pure price competition with left-digit biased consumers also leads to missing prices as shown in Appendix A.2, but I did not solve a full model that can be taken to the data. Insofar as elasticity is a good approximation for the strength of competition it does reduce the ranges of missing prices. But, as I argued above, high elasticity by itself is not a plausible resolution to the discrepancy. Next, consider the case of a multi-product monopolist, as retailers are. Keeping same-price elasticity and left-digit bias fixed, if differentiated products are substitutes, it leads to *higher* next-lowest prices. This result holds whether the other product is more or less elastic, and whether its cost is higher or lower (see Appendix A.3). The conclusion, in any case, is that adding cross elasticities to the counterfactual analysis is likely to make the discrepancy between predicted and observed retailer behavior even stronger.

Another explanation is that the bias level, θ , might be endogenous to the price distribution. Meaning, if all prices were to end at 99 as the estimation suggests, then the bias would have become much lower. This is a sensible argument, but the needed effects on left-digit bias seem unlikely strong. For example, assume that left-digit bias is decreasing in the next-lowest price. Say that it is 0.15 if the next-lowest price is \$2.19. In order for \$2.19 to be an optimal next-lowest price, left-digit bias should decrease to the degree that is eliminated completely if the lowest price used is instead \$2.21⁴¹. This seems like an unlikely strong effect.

Another avenue is that the model is a reasonable approximation, but firms are wrong. One possibility is that firms are following a heuristic (“use 99-prices”) and so when the “absent-bias” price they come up with is close to 99 they round to 99⁴². The appeal of this story is that it makes intuitive sense that an absent-bias price of \$2.00 should be adjusted to \$1.99, but that it is less clear what a \$2.50 should change to. But there is more to it, since not only is it more intuitive to make small changes in prices, but also the impact on profits is larger for the small changes. For example, consider two products with unit costs of \$1 and \$1.25, price elasticity of -2, and bias of 0.15. Absent the bias the product should be priced at \$2 and \$2.50, but with the bias both should be priced at \$1.99. The gains from changing from a naive \$2 to an optimal \$1.99 are 15.5 percentage points in profits, while the gains from changing a naive \$2.50 to an optimal \$1.99 are a much lower, at 2.6 percentage points. So, given the counter-intuitive nature of big adjustments in price, and considering the relatively lower gains of big adjustments (masked in fluctuations in demand in real data), it seems reasonable that firms would stop shy of full adjustment. Making the full adjustment requires either “brave” and large scale experimentation with prices, or a rigorous analysis

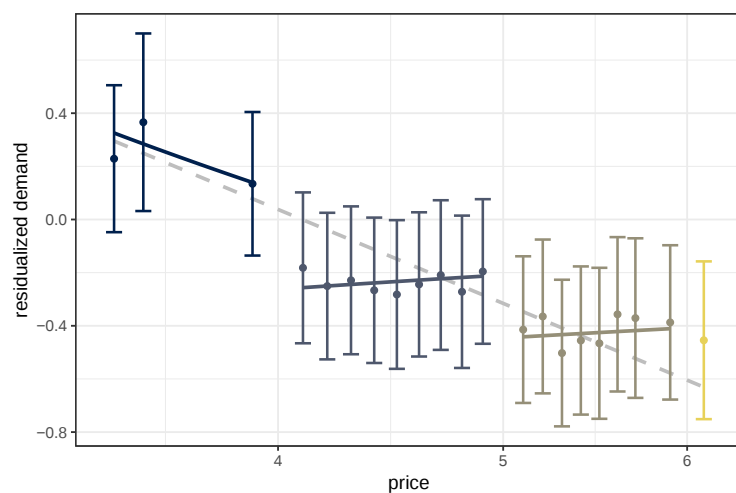
⁴¹To see that, note that if 2.19 is indeed the next-lowest price, it should satisfy optimality such that $\frac{\partial \Pi}{\partial P} = 0$, where I allow for θ to be a function of P . So, it is easy to see that the slope of bias to next-lowest price in that case is $\frac{\partial \theta}{\partial P} = \left(\epsilon (1 - \theta) + \frac{1}{P - P_1} \right) P - \theta$, which equals 8.6 for elasticity of -1.5, bias of 0.15 and P of 2.19.

⁴²also “corroborated” by three anecdotal conversations I held with two mom-and-pop shop owners and a CEO of a large chain of convenience stores.

with a quantitative model.

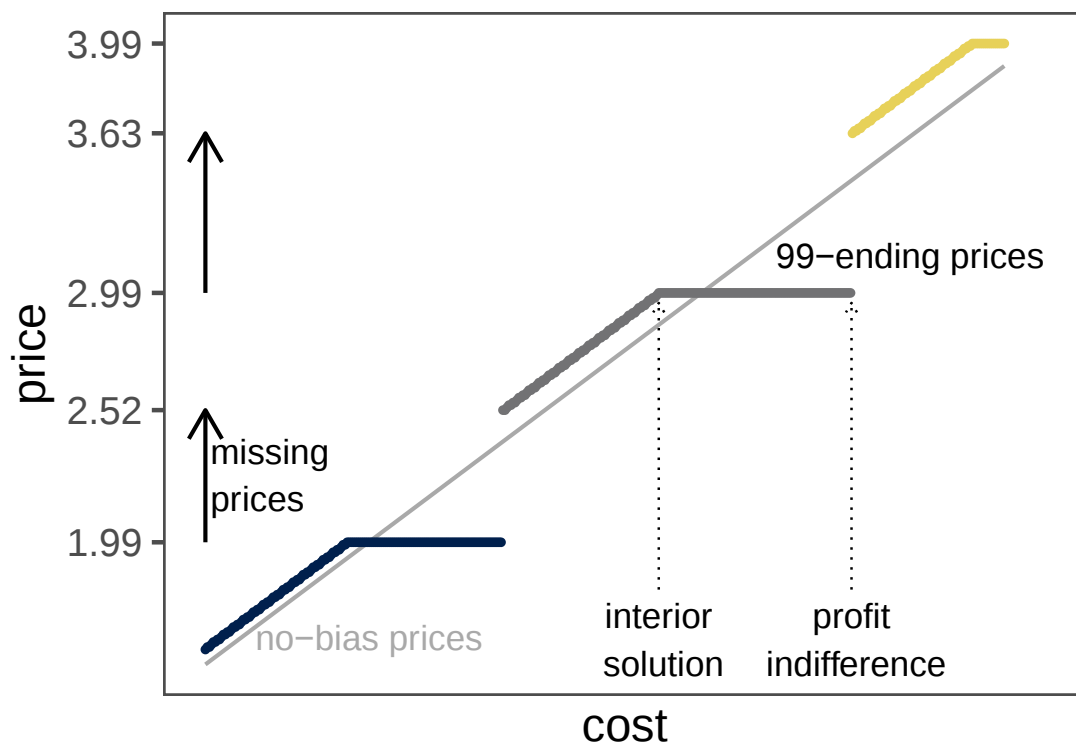
A companion paper, Strulov-Shlain (2019a), analyzes a policy reform in Israel where supermarkets had to stop the common practice of pricing at 99-ending prices, and were only allowed to price “to-the-dime” (i.e, in 0-ending prices). Before the reform about 40%-50% of prices ended with 99, suggesting substantial levels of perceived and actual bias. Using the same model as in this paper, I show that for any reasonable level of left-digit bias, firms should have never round to a 00-ending price. However, all chains did so, and priced about 20% of products at 00-ending products after the reform. Within 6-12 months, they stopped. This finding is inline with the heuristic explanation, showing that firms respond to the bias in a limited “trial and error” sense, but without understanding the underlying model.

Figure 1.1: Demand curve of a single product, exhibiting drops at dollar digits



The figure shows binned demand for a single product (a single UPC). Each dot is the estimate from regressing log-quantity sold on a 10-cent dummy (e.g., \$4.90-\$4.99), controlling for store, seasonality, price of other same-category products, and a promotion fixed effects. To focus on the price support of most common prices, the figure includes all price points with at least 50 observations. Further details in Section 1.4. The item in this figure was not chosen at random, but rather as a motivating example. A systematic analysis is in Section 1.4.

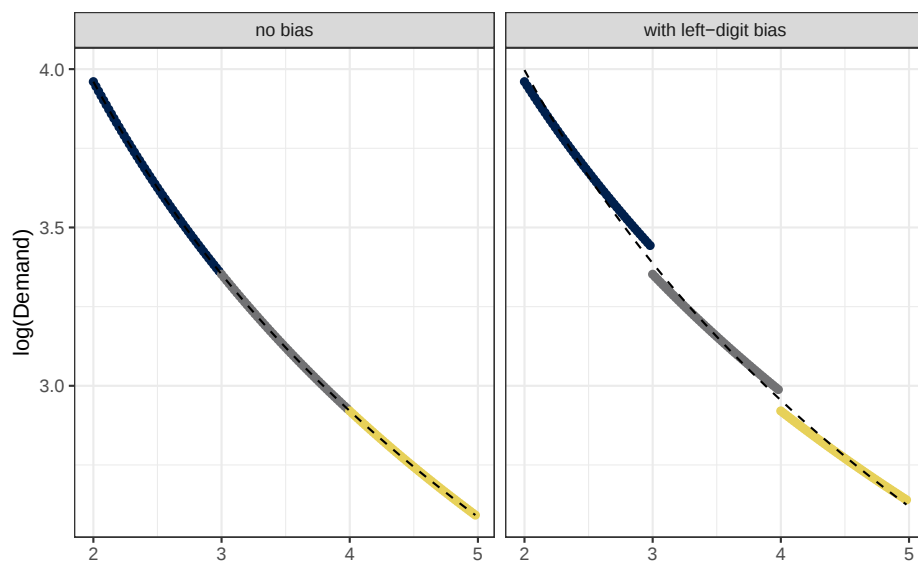
Figure 1.2: Illustration of the optimal price schedule



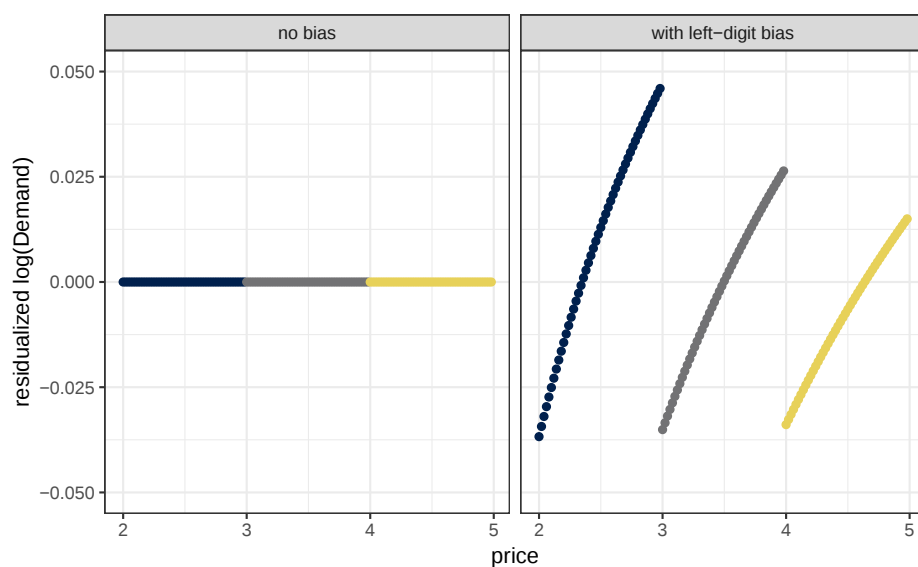
The thick lines represent optimal prices as function of cost, of a monopolist facing constant elasticity demand with left-digit biased consumers. The gray thin line represents the no-bias counterfactual where $p = c \frac{\epsilon}{\epsilon+1}$. The horizontal sections of the thick lines are the 99-ending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

Figure 1.3: Illustration: Demand and residuals with and without left-digit bias

(a) Illustration: Demand with left-digit bias



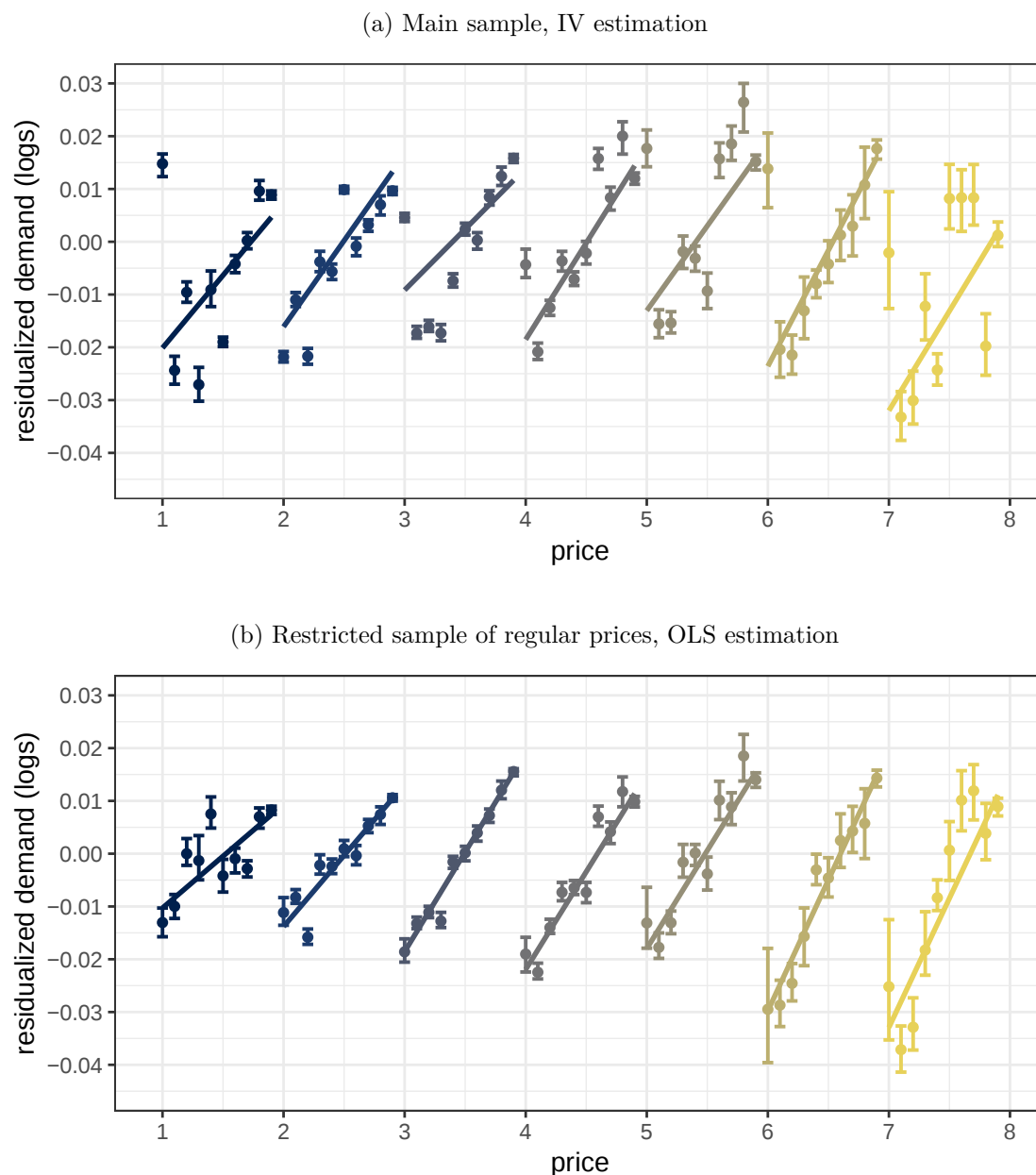
(b) Illustration: Residualized demand from log-log regression with left-digit-biased demand



The figures illustrate sim-

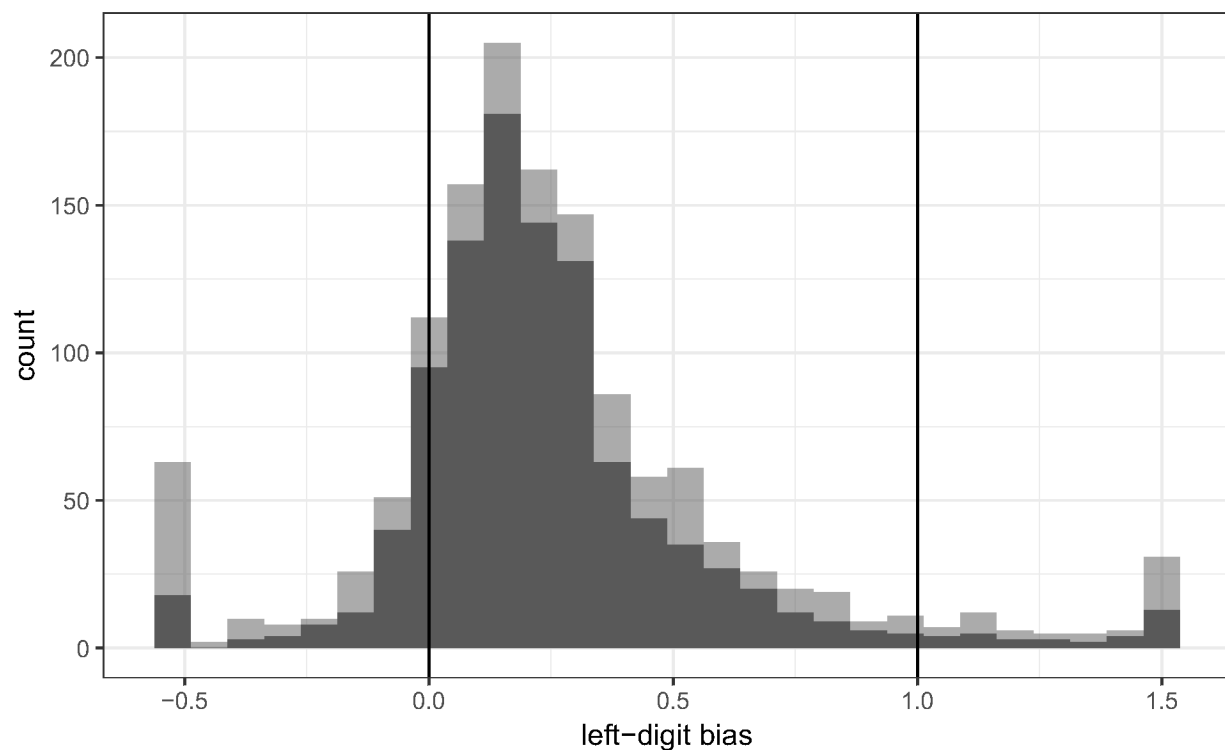
ulated demand curves under a model of left-digit bias. The top figures show simulated $\log(\text{quantity})$ -price demand curves. Colors represent different dollar digits. The dashed gray line is the curve of best fit from a $\log Q \sim \log P$ regression. Bottom figures show the residualized demand (actual minus predicted) by price. Right figures are for a case left-digit bias of $\theta = 0.16$, and left figures are for the standard case of no bias, $\theta = 0$.

Figure 1.4: Sawtooth patterns of residualized demand by price



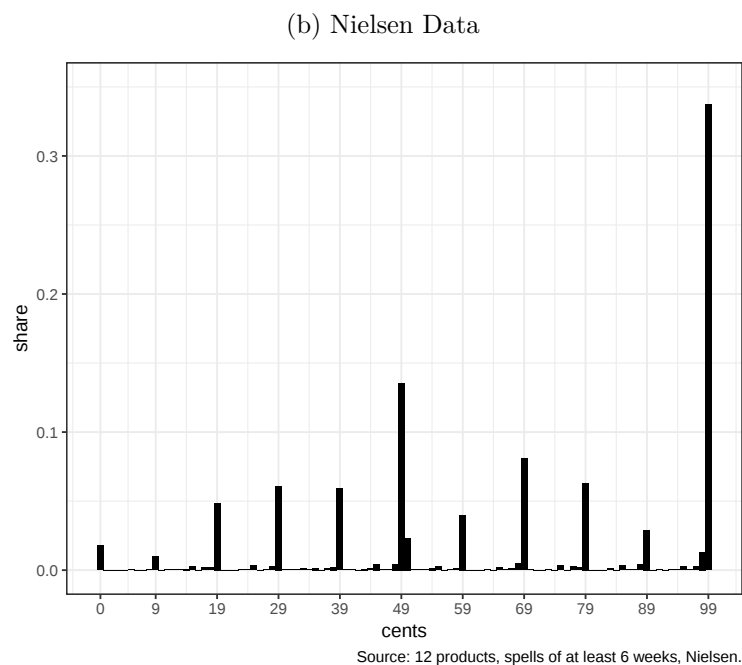
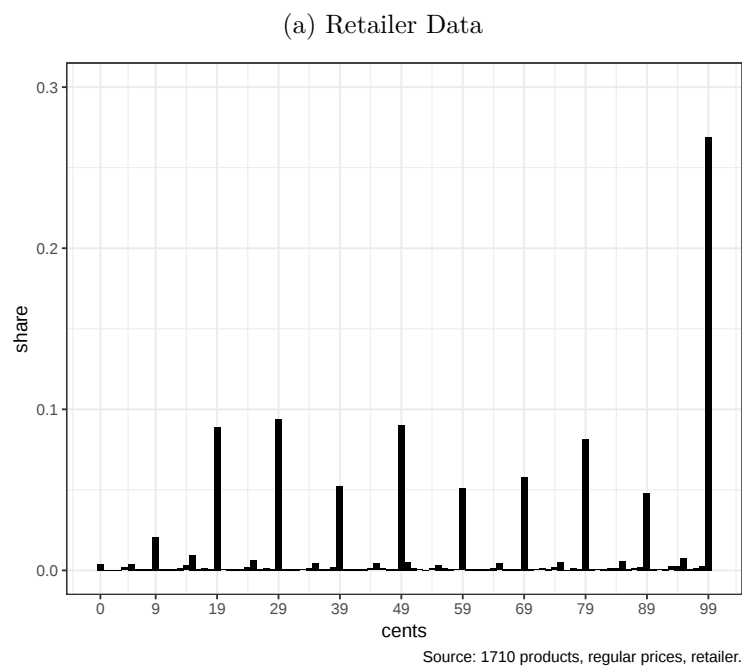
Net log-demand by price, residualizing product-level price elasticity, seasonality, cross elasticities, store fixed effects, and promotion effects. Regressions are conducted at the product level (1710 products, 248 stores, 177 weeks). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Plotted fixed effects estimated at 10-cent bins (e.g. \$3.90-\$3.99), and 95% confidence intervals are the result of 300 cluster-bootstraps by store. The top panel is using data on all to-the-cent prices (Column (3) in Table 1.1), and the regression is run with instrumental variables, as specified in equation 1.6. Bottom panel presents the results of an OLS constructed on regular prices only (column (4) of Table 1.1). Results shown for prices between \$1 and \$7.99 for clarity. See Section 1.4.

Figure 1.5: Histogram of product-level left-digit bias estimates from drops in demand at the dollar-digit



The figure shows the histogram of product-level estimates of left-digit bias, calculated from Equation 1.7, using estimated drops in demand and price elasticity estimated with instrumental variables. Highlighted by darker color are the 75% most precise estimates. Most imprecise estimates are also outliers and vice versa.

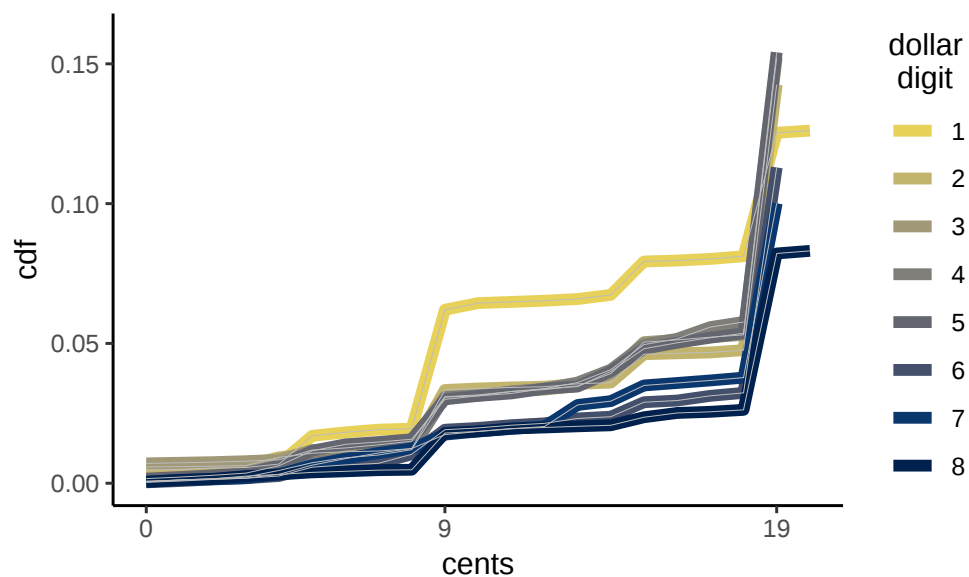
Figure 1.6: Pricing response: Price-endings distributions



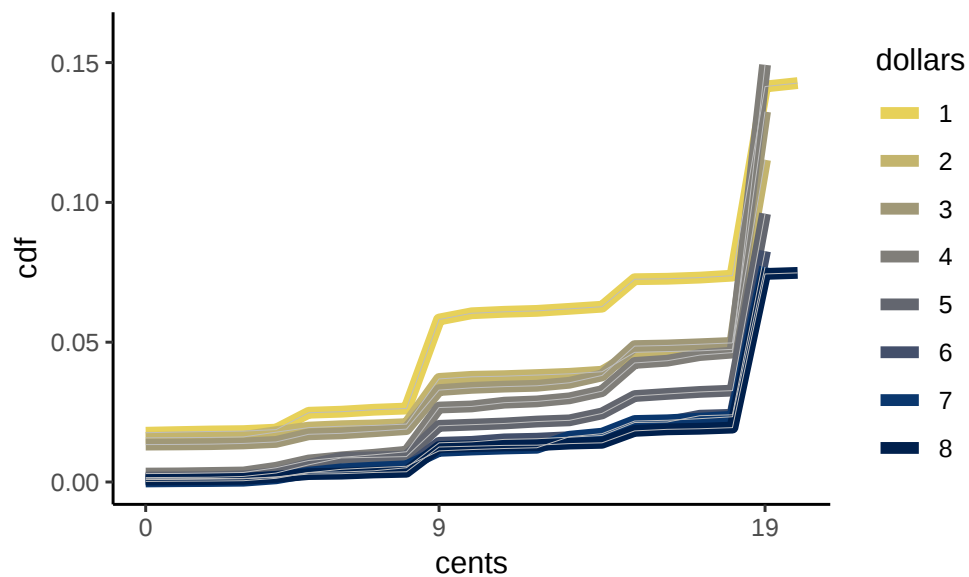
The figures show the price-endings histograms of all regular prices in the retailer data (36M observations, top panel), and all prices that belong to spells of at least 6 weeks in the Nielsen data across more than 60 retailers (11M observations, bottom panel). Main patterns are excess mass at 9- and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

Figure 1.7: Price endings: left-tail of CDF by dollar digit in retailer data

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

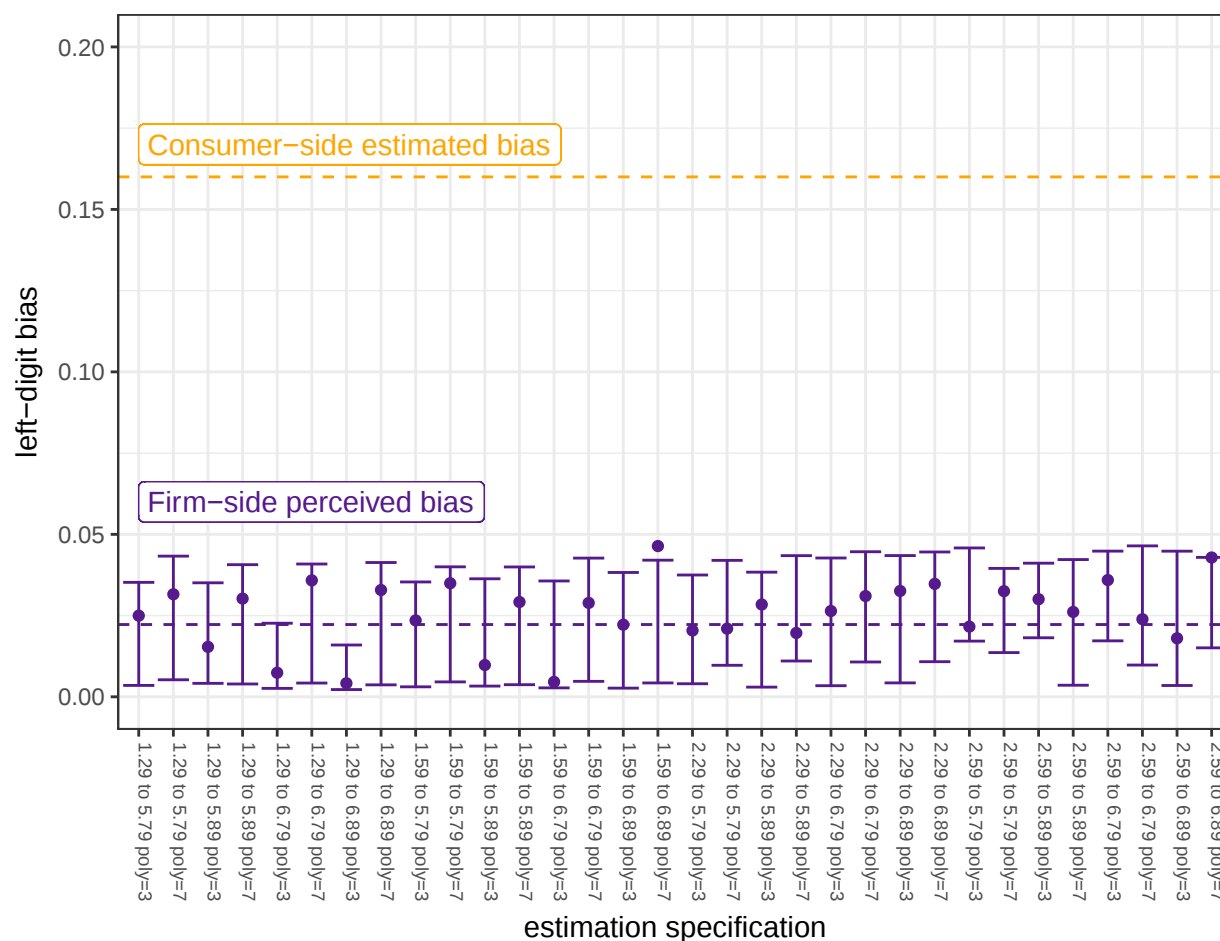


(b) Left tail of CDF for regular prices of full sample



The figures show the left-tails of regular prices price-endings empirical CDFs, by dollar digit. The top panel is the main sample, while the bottom uses all regular prices in the data. Line color represents the first digit of the price, where \$1+ is the lightest and \$8+ is the darkest. The figures show that as prices increase, the firm is using less of lower price endings, qualitatively following the predictions from Corollary 1.2.

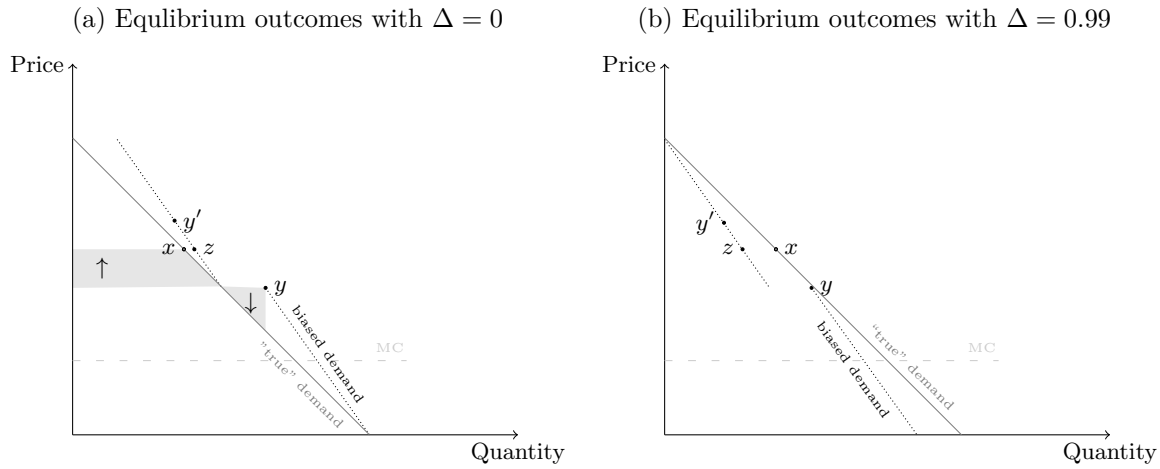
Figure 1.8: Estimated firm-perceived left-digit bias



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

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

Figure 1.9: Illustration of welfare effects



The figures describe the optional equilibrium quantities and prices under different levels of left-digit-bias and firm perception of the bias. Point x is the solution in the standard no-bias case, on the gray no-bias demand line. The no-bias demand line captures the true underlying valuation of the product. In contrast, the dotted line shows the demand under the perceived price $\hat{p} = (1 - \theta)p + \theta(\lfloor p \rfloor + \Delta)$. Panel (a) shows the biased demand when the focal price ending is 0, hence demand is higher everywhere; while panel (b) shows the biased demand when the focal price ending is 99. Points y and y' describe the new possible equilibrium outcomes when there is left-digit bias $\theta = \hat{\theta} > 0$. Point y is the outcome when the optimal price is the lower 99-ending price, and point y' is the outcome for interior price updating. Finally, point z is the outcome when there is bias $\theta > 0$ but the firm prices as if there is no bias, $\hat{\theta} = 0$.

Consumer surplus is the area between true demand and the price for all quantity sold. If there is over-consumption (as in point y of panel (a)), where the price is higher than true demand, this is negative consumer surplus. Producer surplus, or firm profits, is the area between the price and marginal cost for all quantity sold. Deadweight loss is the area between true demand and marginal cost for all unsold quantity in which demand is higher than cost.

Table 1.1: Summary Statistics and Data Selection

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

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

Table 1.2: Estimates of product-level drops and left-digit bias parameter

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

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

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

Panel B: Row (4) provides the estimate from a homogeneous log-linear model, with a single parameter of left-digit bias and single semi-elasticity. Rows (5)-(7) show the product-by-product estimates, inferred by three different approaches: The sufficient statistic approach (Equation 1.7 in Section 1.4); estimated by non-linear least squares (Equation 1.8); and log-linear OLS (Equation 1.9), respectively. To calculate means I winsorize at $\theta = -0.5$ and 1.5. See Figure A.8 for the CDFs of the complete distributions.

Table 1.3: Estimated perceived parameters from firm pricing behavior

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

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

Table 1.4: Profits lost when the firm underestimates left-digit bias

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

Calculated percentage of gross profits lost, relative to optimal pricing. Optimal means correctly perceived bias, $\hat{\theta} = \theta$, while other cases take a lower perceived bias into account. Panel A shows the profits lost if the firm ignores the bias, and panel B shows the profits lost when it underestimates it. Each row makes a different assumption about the “true” parameters and the perceived ones (but always assuming that the firm knows the true price elasticity). For example, under point estimates of left-digit bias of 0.16 and perceived bias of 0.004, and elasticity of -1.5, 2.7 percentage points are lost relative to optimal pricing. If the firm were to ignore the bias altogether, it would lose 4.3 percentage points. The first row in each panel shows the lost profits under point estimates taken from the main analysis. The second row, extreme estimates, makes a firm-forthcoming assumption with lower bias, high price elasticity, and the upper bound of perceived bias. The next three rows explore the role of heterogeneity and how consequential it might be: heterogeneity in left-digit bias, in perceived left-digit bias (truncating from below at 0.001), and in same price elasticity. Finally, the last row takes cross-elasticities into account. At each cell I simulate 3000 costs from a cost distribution that mimics the actual distribution inferred from a constant markup prices. For the heterogeneous cases I draw 50 values for each parameter. For the last row, I sample 80 prices for each of 2 products (from the same cost distribution as above).

Chapter 2

Firm Sophistication: A Natural Experiment on Price Endings

2.1 Introduction

There is ample evidence that firms are better at targeting their consumer biases and limitations than consumers are able to do so themselves (e.g. Akerlof and Shiller, 2015; DellaVigna and Malmendier, 2006; Grubb and Osborne, 2015). “Sophisticated profit-maximizing firms” is a building block of standard economics for sure, but even of behavioral IO (Heidhues and Köszegi, 2018; Spiegler, 2011). But do firms understand their consumers or merely respond to current market conditions, operating with a set of policy functions? This matters for market design, policy design and analysis, and for basic understanding of how firms and the market work. To illustrate with an example, consider a pool player. They need not know the laws of physics to learn how to play from experience. However, now say you make the table round instead of rectangular. If the player understands the laws of physics, they can immediately adjust. But if not, they need to effectively “relearn” these laws again. The length of time it takes to relearn determines if this relearning process has any implications.

Yet, testing the sophistication level of firms is challenging. First, the researcher needs to know how consumers behave and respond to firms actions. Second, the researcher needs to know how should firms respond to the consumers they are facing. Only then we can ask if firms indeed act in a consistent way. If so, it implies that firms are able to learn from experience, but not if they understand the underlying model. To test firms understanding of the model we need a change in the market setting in a way that does not hinder the underlying forces, but that should lead to different actions of firms.

In this paper I propose to study this question in the domain of the empirically established left-digit bias (Lacetera et al., 2012; Strulov-Shlain, 2019b), coupled with a policy reform that altered the admissible prices firms can set. Importantly, this is a domain in which firms, Israeli supermarket chains, exploited the bias before the reform, and were able to keep exploiting it after the reform. That is, this is a setting in which the researcher has testable

predictions, the firm behavior is in-line with these predictions, but then a shock occurs to the market in a way that shouldn't interact with the firms ability to continue to respond.

This paper provides the first empirical evidence for firms' level of understanding of their consumers biases. Firms indeed respond to the bias before and long after the reform. But their prolonged use of dominated prices, in the first year following the reform, points to a lack of an underlying model of consumers behavior and is more inline with heuristic price setting or shallow learning.

The setting and reform that I focus on, are supermarket pricing in Israel from 2013-2015. Israel abolished the 1-Agora and 5-Agorot coins (100 Agorot = 1 Shekel) before 2008, leaving 10-Agorot as the lowest denomination coin. But firms could still price "to the cent". That is, 4.99 was permitted (and prices are already VAT inclusive). If consumers paid with a card, they paid the exact amount; if they paid with cash, the price would be rounded to the nearest "dime" (e.g. 5.00 in that example, or a 4.94 would round to 4.90). On October 17 2013, the government issued a statement that starting January 1st 2014, it will enforce the existing law banning such prices. All prices, starting 2014, have to be "to the dime". Meaning, the modal price ending of 99 is prohibited. In fact, more than 85% of prices had a non-0 (non-dime) price ending and had to change.

A model of left-digit bias, in which prices are perceived with extra weight to the left-most digit was suggested in the literature as explaining the prevalence of 99-ending prices (see Stiving and Winer, 1997 for review). In its simplest form, the perceived price $\hat{p}$ (of a single digit price) is a function of the true price p and left-digit bias parameter θ where

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

Meaning, the perceived price is a mix between the true price and just the left digit with a weight θ . As I show in a companion paper (Strulov-Shlain, 2019b), when consumers perceive the price in such a way, not only do prices bunch below the round number thresholds, but also regions of low-ending prices are now dominated due to the discontinuities in perceived price. Indeed, the model is supported by US supermarket scanner data for a variety of products and chains.

But what should happen when 99, previously a profit-maximizing price ending for many costs and elasticities, is now banned and the just-below price ending can be 90? In this new world there is a larger revenue loss from pricing below the round price threshold (e.g. at 4.90 vs 4.99). Still, I calculate that for almost all reasonable left-digit bias and price elasticities, prices should not end with 00. For example, a previously optimal 4.99 should update after the reform possibly to 4.90, possibly to 5.30, but almost never to 5.00. Therefore, the share of 00-ending prices is a reduced-form measure of misoptimization of firms.

I obtained data on prices of staple supermarket items in stores from three sources. The first two datasets are a panel of prices displayed in stores for specific products as collected by surveyors. One includes about 20 items (e.g. brand X 200gr chocolate bar) with store and chain identifiers, while the other has prices of 170 item types (e.g. chocolate spread) without chain identifiers. The third dataset is a sample of daily revenue and quantities as

collected by registers for 6 items in 295 stores and 12 chains. That is, I have rich price data, and a sample of price-quantity data.

A high share of prices ended with 99 before the policy change (40%-45%), a high share ended with 90 a year or two post-reform (50%-60%). Importantly, there were little 00 price endings before (1.5%) and long after the reform (0%-8%). These results are in line with the model's predictions. In contrast, early post-reform shares of 00-ending prices are quite high, at roughly 20%. The scanner data provides suggestive evidence, though non-significant, that updating a price of the same product in the same store from a 99-ending price to the nearest round price (e.g. 4.99 to 5.00) instead of updating it to a 90-ending price (e.g. 4.99 to 4.90) is accompanied with a few percents drop in demand.

Assuming optimality of 99 for a product in the pre-period and locally constant elasticity, this is a suboptimal transition. Indeed, the products that changed from 99 to 00 tend to later change into 90, while those that change into 90 tend to stay at 90. For some product-stores, there is switching back and forth between 90 and 00, eventually ending at 90. These dynamics show both experimentation and realization of the proper and improper price ending. Also, firms that lower the share of 00 do so in a step-wise fashion and not gradually. The sudden drops reject an alternative explanation in which chains knew they should update to 90 and not 00 but were bounded by some frictions at the product or store level. In addition, a quarter of products change their price every two weeks, so frictions are less likely to be an explanation. I explore the effects of competitive forces, explicitly spatial density of other stores. Evidence is suggestive of quicker convergence for stores that have more nearby competitors to compete with, react to, or imitate.

Meaning, while supermarket chains respond to the left-digit bias by shying away from round prices before the reform, they do not respond in a consistent way after a reform occurs. That is, for a substantial period of time they “unlearned” what they seemed to know before - that round prices are dominated. Further, chains seem to lose a few percents of revenue from these mistakes. I will close with a quote from a CEO of one of the largest supermarket chains, Mega, given shortly after the reform (March 2014):

“There was a 2%-3% decrease in revenues on Jan-Feb ... No one can explain why that happens ...I'm in business for 25 years, and had never seen such months”

2.2 Literature

“Psychological prices” is a term used to describe prices that end with 9, or contain 9 (Foxall, Goldsmith, and Brown, 1998). Its prevalence is a well documented phenomenon (e.g. Levy et al., 2011), and well known to anyone who ever visited a store. Previous studies on the reasons why firms use psychological prices reached mixed conclusions about their profitability (Basu, 1997; Gedenk and Sattler, 1999, see Stiving and Winer, 1997 for a brief review) and suggested alternative explanations for their existence (e.g. Anderson and Simester, 2003; Basu, 1997, 2006; Bizer and Schindler, 2005). Some experimental and empirical analyses

were conducted to see whether and how 9 and 99 price endings affect demand in practice, and found mixed results as well (Bizer and Schindler, 2005; Thomas and Morwitz, 2005).

In a companion paper, Strulov-Shlain, 2019b, I find large scale evidence supporting a specific explanation in US supermarkets. Prices, and especially regular (non-sale) prices, are set just-below round numbers because demand drops by a few percents when the dollar digit changes.

Whatever the explanation may be, Gedenk and Sattler, 1999 show that under a variety of demand structures, it is more profitable for firms to err on the side of consumers having discontinuous drop in demand when crossing left-digit thresholds. In fact, they claim that for reasonable elasticities, even if consumers drop-off the last digits as little as 3% of the time, it is profitable for firms to end prices at 9. Indeed, Strulov-Shlain, 2019b finds similar implications, but there I highlight the dominated low-ending prices.

Outside the domain of prices in stores, some recent papers have studied closely related models of inattention (see review by Gabaix, 2019) and left-digit bias (Busse et al., 2013; Lacetera et al., 2012). These models are smoother in the sense that inattention is parameterized as a number between 0 and 1 (call it θ), interpreted as either the share of inattentive consumers or some individual level stochastic inattention. Leading empirical findings are of Chetty et al., 2009, who think of θ as a measure of saliency of sales taxes; Lacetera et al., 2012 who show the downward pricing curve of used cars by their odometer reading, with discrete downward jumps when the 10k and 100k digits change; and Hossain and Morgan, 2006 who concentrate on people ignoring shipping costs when purchasing from eBay. The inattention measures in these papers are quite high, in the range of 20%-66%.

Prior work in behavioral industrial organization has primarily focused on biases from the consumer's standpoint, and how perfectly-knowledgable firms can exploit those biases in market equilibrium, both theoretically (such as in DellaVigna and Malmendier, 2004; Gabaix and Laibson, 2009, and the book of Spiegler, 2011) and empirically, in counterfactuals (such as in Agarwal, Chomsisengphet, Mahoney, and Stroebel, 2015; Grubb and Osborne, 2015). Comparatively, far less work details the extent to which firms do in fact succeed at exploiting such biases. One barrier is to find a natural experiments that will allow a window to observe these dynamic responses. A challenge that this paper aims to grapple with.

2.3 Model

I am using the model from Strulov-Shlain, 2019b, where biased consumers meet monopolistic firms. In the model, a price is perceived in a distorted way, as a mix with weight θ on a focal price ending Δ , and $1 - \theta$ on the true price. It causes the perceived price to change discontinuously when the left-most digit changes, and also to be less sensitive to price changes that do not involve left-most digit changes. For example, with $\theta = 0.2$ a change between 4.99 and 5.00 is perceived as a 20 cents increase, while a change from 5.00 to 5.01 is only an 80 cent increase. Formally,

$$\hat{p} = \hat{p}(p; \theta) = (1 - \theta)p + \theta(\lfloor p \rfloor + \Delta) \quad (2.1)$$

These perceived prices then affect demand. Demand is a function of the perceived price, while consumers still pay the true price when purchasing an item. Assume a demand curve with constant elasticity, which in this setting should only hold locally, and set $\Delta = 0$ for simplicity. The demand curve then is

$$D(p; \theta) = A\hat{p}^\epsilon = A((1 - \theta)p + \theta \lfloor p \rfloor)^\epsilon \quad (2.2)$$

With accompanied profits of

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

where the assumption of a unit cost c is an approximation for supermarket pricing. It can be shown that prices under this demand structure bunch just below round numbers, with ranges of missing prices pinned down by the parameters of the problem.

Definition 1. *The Next-Lowest Price P is the lowest price used above a just-below price q , solving the following implicit equation:*

$$P + \frac{(1 - \theta)P + \theta P_1}{\epsilon(1 - \theta)} - \frac{((1 - \theta)P + \theta P_1)^\epsilon P - ((1 - \theta)q + \theta q_1)^\epsilon q}{((1 - \theta)P + \theta P_1)^\epsilon - ((1 - \theta)q + \theta q_1)^\epsilon} = 0 \quad (2.3)$$

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

For what parameter levels may we observe 00-ending prices? Whenever the next-lower price $P > P_1$ we will not. Since for any elasticity, if $\theta = 0$ there is no discontinuity in demand, there must be some $\theta_0(\epsilon)$ for which $P = P_1$. Then, since stronger left-digit bias θ increases P (see Proposition 2 in Strulov-Shlain, 2019b), for any $\theta > \theta_0$, $P > P_1$ ¹. Regrettably, there does not seem to be an analytical solution for θ_0 . Solving for it numerically we get that for an extreme elasticity of -7 and $P_1 = 2$ (and from Proposition 2 in Strulov-Shlain, 2019b less elastic demand increases P , and by Corollary 1, P is increasing in P_1 , θ_0 is a mere 0.00075, we will observe no .00 pricing when .99 is available. In contrast, I find that average left-digit bias is $\theta = 0.16$ and typical elasticities in these settings are about -1.5 to -4 (DellaVigna and Gentzkow, 2017; Hausman et al., 1994; Nevo, 2001).

However, the reform made the trade-off starker. Now the just-below price is 90-ending and not 99-ending. That is, products that ended with 99 before, such as 2.99, had to be repriced, for example at 2.90, 3.00 or 3.10. Perhaps with this stronger trade-off, 00-ending prices are now optimal? We can answer that question in the same way as above, asking for what parameters would $P_{0.1} > 0$ given that the just-below price is now $q = (q_1, 0.9, 0)$.

¹since prices are actually discrete, I will ask for what θ_0 we get that $P = (P_1, 0, 0.01)$. Meaning, that the next lowest price is 1 cent above the round price.

Of course, if $\theta = 0$ and costs are well distributed, there will be prices with $P_{0.1} = 0$, so the question is how large should the bias be (even for large elasticities such as 5, and low prices such as 2.00) to prevent this. Two forces are making the trade-off more favorable to .00 ending prices. First, the just-below price is now 90-ending instead of 99-ending, which leads to lower profits if the optimal price is 99-ending. Second, the next price above 00 is not 01-ending but 10-ending. These two forces do have a combined strong effect on the minimal bias leading to no 00-prices, but the levels are still very low. The threshold levels are depicted in Figure 2.1. Indeed, while before the reform, with elasticity of -4 and $P_1 = 2$, the minimal bias is practically zero, after the reform the minimal bias becomes $\theta_0^{Post} = 0.02$, which is still far from 0.16 (levels found in US scanner data), but not outside the scope of what firms might consider the bias to be.

Figure 2.1 shows the threshold values of $\theta_0(\epsilon)$ for different first-digits, before and after the reform. The dark lines are the pre-reform thresholds, showing that 00-ending prices should not be observed for almost any value. The light curves are for a post-reform setting, showing that the threshold indeed increase substantially, but not enough to justify 00-ending pricing given common estimates of θ .

The above results in the following prediction:

Conjecture 1. *Under levels of left-digit bias and local price elasticities that are common in the data, pricing at 00-ending prices is not profit maximizing. Therefore, after the reform, firms should avoid 00-ending prices.*

2.4 Data and Setting

On October 17th 2013, the Ministry of Economics announced that starting January 1st 2014 it will start to enforce an existing law banning pricing in non-existing coins². Since Israel abolished the 1-Agora and 5-Agorot coins (The Shekel is consisted of 100 Agorot) in 1991 and 2008 respectively, and since many prices ended in non 10-Agorot (“dime”), most prices had to change. This policy came about since consumers felt they were being tricked by non-existing prices. Before 2014, the consumer paid the posted price (prices are VAT inclusive) when paying with a card or the price rounded to the nearest dime if paying in cash.

I have two data sources following prices before and after the policy change, both are panel data where the unit of observation is a product in a store. The main sample was collected by a consumers advocacy group, “Israel Consumer Council” (ICC). Their employees went to supermarket stores with a varying list of products every 2-4 weeks between early 2013 to early 2015 and wrote down the prices of the goods on the list. A product can be thought of as a UPC, of unique producer, weight and size. ICC collected the data to later post it on their website and inform consumers which store in their surroundings sells the cheapest bundle that week. The big advantage of the data is that it includes an approximate location and

²except for “continuous” products - gas, water, and electricity

store name. These variables allow to analyze heterogeneity between and within supermarket chains.

The data is not a balanced panel, since the product list is changing over time, stores enter and exit the sample, and the sampling rate is inconsistent. Although there does not seem to be strong selection, to overcome these issues I clean the data to be more balanced. First, I restrict the sample to products that are sampled at all periods. Second, I restrict it to stores that are sampled at least T times within each policy-relevant period³, and at the last period in the data (on February 2015). Third, I balance the panel by completing missing price observations for a product in a store backwards⁴, i.e. if product's prices in 3 consecutive sampling periods are $\{3.9, \text{Missing}, 4.2\}$ I complete them to be $\{3.9, 4.2, 4.2\}$. Table 2.1 shows the differences in the number of products, stores and prices between the samples. The main sample consists of 21 products in 109 stores. Importantly, more expensive products were not sampled at 2015 (e.g., wine, feta cheese, corn schnitzel) which creates a big difference in the mean price. The data cleaning indeed make the patterns cleaner, and is crucial for exploiting the panel nature of the data for analysis, but does not qualitatively alter the main results. An important trait is that prices are recorded for the first unit purchased. That is, if there is a promotion such as "2nd item for 50% off" this is not taken into account. As such, the prices over-represent non-sale prices.

The second database are prices collected by Israel's Central Bureau of Statistics (CBS) in order to create the Consumer Price Index. The data are similar in nature, in the sense that surveyors collect displayed prices of a list of products from stores every month. Some important differences between the datasets exist, since the goal of the CBS is to create a reflection of the representative shopping list. Therefore, the CBS collects an extensive share of data from non-supermarkets (e.g., markets, specialty shops and convenience stores); each product is sampled from a somewhat different set of stores; and products are a "product-type" rather than a UPC. For example, the product "cottage cheese" may be of different size, manufacturer and fat content between stores and within stores over time, and collected mainly from supermarkets in big cities while "Zaatar Spice" will be collected mainly in markets and in the Northern and Southern rural areas. However, the panel is balanced, sampled monthly between 1/2012 to 12/2015, and the set of product-types is an order of magnitude larger than in the ICC data (with 171 product-types in the final sample). Summary statistics can also be seen in Table 2.1. To make the data more comparable to the ICC data we restrict it in the same way, and consider supermarkets only.

Most of the analysis is done using the ICC data, since it allows for analyzing the heterogeneity between chains which is of main interest. The basic analysis is conducted for both the ICC and CBS data, showing very similar results.

Finally, a third dataset is daily scanner data provided by StoreNext Ltd, a market research firm collecting transaction level data from store registers and aggregating it. The unit of

³specifically, we chose 4,3 and 5 times for the pre-announcement (3/13-10/13), between announcement and enactment (10/13-12/13), and post enactment periods (1/14-2/15) respectively for the Main ICC data; and 15 times at the post period for the Post-only analysis.

⁴we also complete forward as a robustness test.

observation is a product (UPC) in a store in a day. I have purchased data for a sample of 10 UPCs from a balanced set of 295 supermarket stores over 3 years (2013-2015). However, one of the UPCs has a single price throughout the sample (canned corn), one other is a product that was not available before the reform (400gr Hummus). Two others are Cottage cheeses that follow uncommon pricing behavior (see Hendel, Lach, and Spiegel, 2017), leaving me with 6 products that exhibited 99 pricing before the reform.

The data have some clear measurement errors in it with unknown source. Regretfully, the data quality issues require extensive cleaning and make the analysis underpowered and results are not robust. To understand the issues, consider the following series of prices in the data, of a chocolate in the same store from four consecutive days in 2015. The price is the division of revenue by units sold in that day. On Sunday the price was 10.9025 (4 units sold), on Monday exactly 10.90 (1 unit sold), and on Tuesday it was 10.902308 (13 units sold). That is, the price on Sunday seems like an average of 3 units sold for 10.90 and 1 unit for 10.91 (and on Tuesday of 10 units for 10.90 and 3 for 10.91). A same-day price difference of 1 Agora is extremely unlikely and 10.91 was not an admissible price in 2015. Of course, when the price is so close to 10.90 it is easy to guess that 10.90 is the true price, but prices also fluctuate by a few Agorot from day to day, which makes it harder to determine. To get a sense of the magnitude of such noise in the data, while weekly level data from a national US retailer exhibit 90% of prices that are “to-the-cent”, the *daily* Israeli data shows only 55% of those.

Given these issues, I round prices to the nearest admissible price (at the 1-Agora level in 2013 and 10-Agorot level in 2014-2015), and only keep prices that are exactly round or are part of a same-price spell of at least 3 days. Two consecutive (rounded) prices are in the same-price spell if they are within 2-Agorot from one another (e.g. 4.89 and 4.87), and the length of the “same-price spell” is the number of successive days of these identical or close prices. I am keeping product-store pairs where a product sold for at least 80% of data days. For the final sample, I am also excluding the 4 products mentioned above, and observations on Saturdays since most chains do not operate on Saturdays.

2.5 Empirical Analysis

Design

As shown above, the main prediction of left-digit bias in consumers demand, is that firms should avoid pricing at round prices. Further, even following the reform, firms should still avoid pricing at round prices. The reasoning is that a change from, say, 4.99 to 5.00, is accompanied by a significant drop in demand. For example, with left-digit bias of 0.1 and price elasticity of -2, demand should drop by about 4.5%.

With that, I ask what are the shares of 00-ending prices before the reform (pre); between the announcement and enactment of it (interim); and after enactment (post).

First, assume the firms are always optimizing. In the standard case, with no left-digit bias, we should observe roughly equal shares of 00-ending prices as other price endings (see El Sehity et al., 2005). For example, equal ratio of 00-ending prices to 99-ending prices before the reform, and equal ratio of 00-ending prices share to 90-ending prices after the reform. In contrast, if there is left-digit bias, the shares of 99-endings pre and 90-ending prices post should be high, and there should be almost no 00-ending prices before or after. Finally, if firms only respond to the bias but do not understand its source, we should see high 99-ending shares and low 00-ending prices pre-reform, high 90-ending prices and low 00-ending prices post-reform, with a possible transition period of high 00-ending prices at the interim and early pre-period.

When considering demand, we would expect that within a smooth demand estimation, there will be a few percents difference in demand between a 99-ending and 00-ending adjacent prices. The ideal experiment is one where prices of the same item in the same store change between 99-ending and adjacent 00-ending randomly and in random order. Firms respond to the reform in what seems like randomly rounding prices to 00-endings. Even with perfect data, the order is not random (always from 99 to 00) which makes it challenging to separate time effects from price-ending effects (still possible if assuming other price endings are as random). However, data is not infinite and not clean enough. Still, considering the gap between 99 and 00-endings can give a sense of the demand drop.

Pricing

Over all products and stores, the share of products that end with 99 before the policy announcement is about 45% and kept stable for both the ICC and CBS data. In contrast, the shares of 00 and 90 were low, with 1.2% of 00 and 1.7% of 90.⁵ Furthermore, the share of 90-endings is at 65% a year after the reform in the ICC data, and at about 52% in the CBS data. That is, a year after the policy change, firms continue to bunch just below the round price threshold. These patterns are consistent with firms responding to left-digit bias.

In section 2.3 that a firm's best response to the reform should be not to update prices from 99 to 00. Figure 2.2 shows a repeated cross-section snapshots of the shares of products that end with 99, 90, and 00 by sampling periods. The shaded area represents the period between announcement and enactment of the policy change. The share of products that end with 99 is high and stable at the pre-period, while the shares of 00 and 90 are very low. Importantly, firms shy away from 00 a year after the policy change. However, they do price extensively at 00 initially.

After the policy is announced, firms start to change prices immediately. From basic principles, a constrained set of actions (prices rounded to 10-Agorot instead of 1-Agora) should be less profitable. Therefore, the immediate and gradual shift suggest one of two explanations. First, that firms face frictions in price changing, and had to start updating prices immediately to manage to abide the new policy in time. Alternatively, that they

⁵the second and third most frequent price endings were 49 (6.6%) and 89 (3%)

wanted to understand the demand with the less-tested price endings before changing all prices.

As the policy starts to bind on January 2014, there are about 42% of products that end with 90 and 20% that end with 00. The high share of 00-endings goes against the model predictions. However, there is a clear downward trend in the share of 00s over time. What drives this downward trend? I conjecture that indeed 00-ending is a suboptimal price, and firms re-learn that fact slowly.

Chain Heterogeneity I argue that firms learn that 00 is a sub optimal price-ending over time, and once they realize it they stop pricing with it. Figure 2.3 shows the shares by chain, for the biggest 6 chains in the data after the policy change⁶. For most chains, the share of 00 behaves in almost a step-wise fashion, where for “Mega in the City” and “Rami Levy” (the second and third most largest chains in the market) the shares go clearly to almost zero. “Shufersal” (the largest chain) with its two subsidiaries “Deal” (discount stores) and “Shelli” (urban stores) going to 10% of 00 price endings in mid 2014. “Wine Pavilion” and “Shufersal Deal” seem to also go to almost no 00s by 2015, but the data ends too soon to see that convincingly⁷.

Can these patterns be affected by frictions? A first kind of friction is a sort of price stickiness friction driven by “menu costs” or other exogenous price updating constraints (e.g. Nakamura and Steinsson, 2008). These seem less relevant here, since 87% of all prices in the pre-period had a non-dime level price ending. Meaning, almost all prices *had* to change due to the reform. Therefore, If firms are aware of the bias, they should change to something that is not 00-ending. A second, more convoluted type of friction is if firms initially make a mistake of updating prices to 00-endings, quickly realize that it is a mistake, but then slowly update due to “regular” price frictions. Though price changes are frequent (in ICC data, roughly 25% prices change biweekly), there might be a staggered effect. In that case we might expect to observe a gradual decrease of the share of 00-ending prices in the post-period. A sharp decline in the shares of 00-ending prices, described in the above paragraph and evident in Figure 2.3, is hence evidence against frictions.

Product Heterogeneity Figures 2.2 and 2.3 aggregate movements across the different products. In Strulov-Shlain, 2019b I estimate substantial heterogeneity of left-digit bias across products, even though in the US data this heterogeneity is not correlated with pricing behavior. The reform allows to ask finer questions. Consider the assumption that products with higher shares of 99-endings are associated with stronger left-digit bias, or at least perceived left-digit bias⁸. The correlation is perfect if different products have the same cost

⁶I changed the sample here to be less sparse in the post period, with product-stores with at least 15 observations (out of 21), to make the patterns cleaner. The usual sample is not as clean, but qualitatively similar.

⁷This test would have been cleaner in the CBS data, however there we do not have firm identity.

⁸What does it mean to have bias heterogeneity between products? One explanation is that different clientele purchase different products, another is that other characteristics of the products receive more

distribution and elasticities. When elasticities differ but costs are smoothly distributed, a positive correlation remains, even if not perfect. Then, stronger left-digit bias creates a stronger penalty of pricing at 00-ending prices. True, even with heterogeneity, 00-ending prices are dominated for low values of left-digit bias. Yet, assume there is some fixed benefit from pricing at 00-ending prices. To the extent that more 99-endings reflect stronger perceived left-digit bias, these should be associated with less 00-ending prices. Meaning, due to true or perceived left-digit bias, if firms understand the model we would expect items that had more 99-endings to have fewer 00-endings post reform. Do these predictions hold in the data?

Figure 2.4 gives an idea about how firms responded. The figures show for each product the share of 00-ending prices at different periods against the share of 99-ending for these products in the pre-period. Starting at the pre-period, Figure 2.4a shows that even for products with very few 99-endings (presumably of lower left-digit bias), there was almost no 00-ending pricing taking place before the reform. A year after the reform, as shown in Figure 2.4c there were more 00-ending prices for some products, and mostly for items with lower shares of 99-endings. Finally, immediately after the reform, as shown in Figure 2.4b, there were much more 00-ending prices, and here too, mostly for items with lower shares of 99-endings though the negative correlation is driven by one group of products⁹. In appendix Figure B.1, we see similar patterns in the CBS data, though less pronounced, where each point is a product-type rather than a specific product.

Figures 2.4 mask the role of chains and individual stores, so I turn to a simple regression analysis. Table 2.2 shows 9 coefficients from separate regressions at the product-store level. For each product-store I calculate the share of 99-endings in the pre-reform period (January to October 2013), and the 00-endings of products in 3 periods - pre-reform, early after the reform (1-2/2014) and a year after the reform (1-2/2015). The findings are consistent with the figures, showing that more 99-shares were not correlated with 00-shares before the reform (even though there a stronger correlation is mechanical), was strong early, and weaker (since also levels decreased overall) a year after the reform.

Learning Method: Price Ending Paths Further insights can be gained from looking at the price-ending paths of products. While we know that the share of 00-endings was initially higher post-reform, and then declined, there are different processes that can generate these aggregate patterns. For example, if stores learn from experimenting then we might expect them to switch between price endings, and then once they realize that 00-endings is a bad price ending stop using it. If in contrast they can compare demand at a price ending to some expected demand, then we might expect some price endings to be absorbing (e.g. 90-endings) and some to be unstable (or “Garden of Eden” states, e.g. 00-endings). An absorbing state

attention than the price (such as observable quality in fresh produce), or that items purchased in bulk or multiple units require price multiplication which puts extra weight on the left-most digits.

⁹One group of products, of fresh produce, was very high on 99-endings pre-reform and translated immediately to 90-endings post-reform. The reason for these items different pricing might rely on that they are purchased by weight (the pricing norm in Israel is to price per kg), but that is only a conjecture.

is one where demand is as good or beyond expectations, while a Garden of Eden state is one where demand is lower than expected, so the retailer changes the price. The following exercise aims at testing these processes.

Consider the price path of a product in a store that was initially priced at 99-ending. Categorize price endings as being 90, 00, or Other. Then, at each period, count how many products changed from 99 to 90, to 00, or to “Other”. If a product changed into 90, then ask what was its price ending at the next period, and so on. This allows to categorize all products into one of several states representing the price-ending path, as described by the automaton (state-machine) in Figure 2.5. The results of the exercise are shown in Figure 2.6, showing the shares of the different price-ending paths as different shades. The green represents 99, the blue shades 90 (or Other after 90,) and the red shades 00 (or Other after 00)¹⁰. First, the absorbing/Garden-of-Eden states are striking. For example, products that initially changed to 90-ending stay at 90 as the dark blue ribbon is constant in size. In contrast, products that initially changed to 00-ending don’t stay at 00-ending (the shrinking “to 0” area). Further, the initial 00-endings later change mostly into 90 (the “to 0 to 90” shaded area). Compare that to the area of the 90-ending after 00-ending (the lighter shades of blue), which are all larger and growing in size. At the final period in the data, these products are more likely to end with 90 than 00 by a ratio of 15:1, and more likely to end with anything else but 90 than 00 by a ratio of 3.5:1. So, there is suggestive evidence consistent with a “learning from mistakes” kind of learning.

Finally, consider the lightest blue and orange. These are the paths of products whose price-ending switched back and forth between 90 and 00 (e.g. $99 \rightarrow 90 \rightarrow 00 \rightarrow 90$). This kind of behavior might suggest that firms are trying to experiment and understand the demand structure. At the very least, these are not pricing decisions driven by frictions. Still, comparing the size of the light blue and orange, the switching products are 8.4 times more likely to land at 90-ending than to stay at 00-ending after experimenting.

Can competitive forces increase the speed of convergence? A partial answer was already given in Figure 2.3, showing that there is large heterogeneity between chains. Meaning, since all chains are national, limited unexplained variance is left for local forces. However, we can still ask whether stores that have more nearby competitors lower the share of 00 faster than those with fewer competitors. For each store in our data, I count the number of competing stores within a radius. For example, above and below median within a 5 km radius¹¹. I then compare the differences in shares between stores above and below the median of spatial competition. Figure 2.7a shows that stores in a denser area have consistently a lower share of 00, at least initially. The dashed lines represent significant differences when controlling also

¹⁰The green are the products that ended with 9. Almost all of these changed at one point or another to either 90 or 00. Rarely is there a product that was priced at 99-ending and changed its price to anything else (In the figure, if changed to Other remains green at the post period).

¹¹Choice of other radii does not seem to matter. A choice of 5km keeps stores above and below median also within more densely populated areas, and is not a simple separation of urban versus sparse villages. I could not get too narrow due to measurement error in locations, which are based on store name and city, only sometimes with a precise address, and then extracted from Google Maps.

for product-level shares, and clustering standard errors at the store level. The differences are significant mainly at the second half of 2014, implying that maybe differential learning is in place. However, different chains are located differently between more or less dense regions, and the between-chain heterogeneity can be the source of differences. Therefore, I do the same exercise only for the 25 stores of “Shufersal Deal” (Figure 2.7b), where differences in shares of 00 above and below the median¹² are less pronounced but still similar qualitatively. Most of the gap is there on January 2014, and is somewhat persistent, suggesting if anything a negative effect on speed of convergence.

Demand

In Strulov-Shlain, 2019b I use the variation within dollar-digits to extrapolate demand and estimate the projected differences in demand between a, say, \$3.99 and \$4.00, because it is rarely the case that a product will be charged both prices. The uniqueness of the Israeli reform and the way firms responded to it allows to compare the demand at these two prices (or between 90-ending and 00-ending prices) because of the post-reform prevalence of 00-ending prices. This is the exercise I am doing here, even though given data limitations as described above, it is only suggestive and results are not robust and insignificant.

I estimate the following specification

$$\log Q_{ist} = \alpha^{90} \kappa_{ist}^{90} + \alpha^{00} \kappa_{ist}^{00} + \alpha^{99} \kappa_{ist}^{99} + \beta_{is} \log P_{ist} + X' \beta + \epsilon_{ist} \quad (2.4)$$

where i is product, s is store, t is date (day). Q are number of units sold and P is the average price. The α are the price-endings fixed effects, such that κ_{ist}^{99} equals 1 if a price ends with 99 on the second half of 2013 (0 otherwise), κ_{ist}^{90} and κ_{ist}^{00} equal 1 if the price ends with 90 or 00, respectively, during the first half of 2014 (0 otherwise). X includes various controls - a product by store fixed effect, same price spell length by product-chain fixed effect (to capture transitory sales effects), a week-of-year by product-chain fixed effect (to capture seasonality), year by product-chain fixed effects, and product-chain by day-of-the-week fixed effect (to capture within week seasonality of traffic).

The most interesting comparison is between α^{90} and α^{00} , because the α^{99} is confounded with the pre-period, which makes it trickier to interpret the effects. For example, if many of the prices tend to end with 99 during 2013, the demand slope will be closer to the 99-ending prices due to the year fixed effects. Yet, removing the year fixed-effect will confound the price-ending effects with other year effects. Comparing the 90-ending and 00-ending, after controlling for the general demand shape, hence is not subject to that critique.

As example of results consider Table 2.3. Results vary by the specification, and the differences between α^{90} and α^{00} are not statistically significant even if they are of the right magnitude. A finer example, comparing the change in demand of items that changed their price from 99-ending price to the nearest 00-ending price within the same store yields results that are unreasonable in magnitude, so are not presented here.

¹²The allocation to medians is still based on all chains

The results are inconclusive, but documented here for the purpose of continuing research in this avenue, with hopefully better quality data.

2.6 Conclusions

Supermarket chains did not seem to have a model of their consumers demand structure, even though they responded to it in the long run by bunching at 99-ending and 90-ending prices and avoiding 00-ending prices. Specifically, in the months following the reform, firms used 00-ending prices extensively, only to stop again later on. The patterns and heterogeneity of these price changes suggest limited response to incentives, and a local form of learning. Demand estimation is inconclusive given available data.

The implications for the behavioral IO literature is that one must be cognizant of how well can firms exploit their consumer biases. The assumption, taken in mechanism design and underlying welfare calculations, that firms are sophisticated and know how to respond perfectly to their market - how to price, what to hide, how to price discriminate or compete - might not pass muster. This gap, between how we model firms behavior and how firms might actually behave, can have implications in designing regulation or evaluating it. For example, say firms take a year to figure out how to overcome bans on some hidden fees. An empirical analysis of the response in the short term might depict a too optimistic picture, while a theoretical/long-term analysis might be too pessimistic. Perhaps recurring changes can alleviate the possibility of firms to take advantage of their consumers biases.

Taken together with the evidence in Strulov-Shlain, 2019b, the theme is of firms that respond partially to left-digit bias by following rules-of-thumb without constructing a model of the bias.

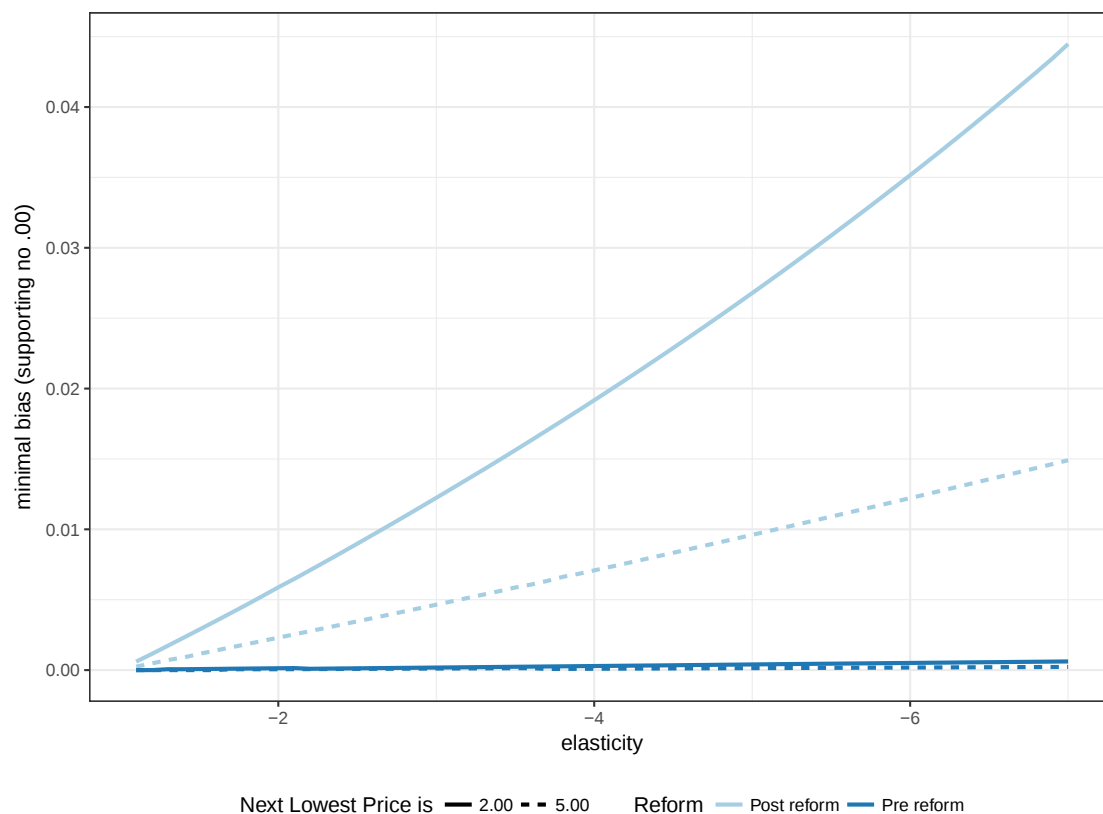
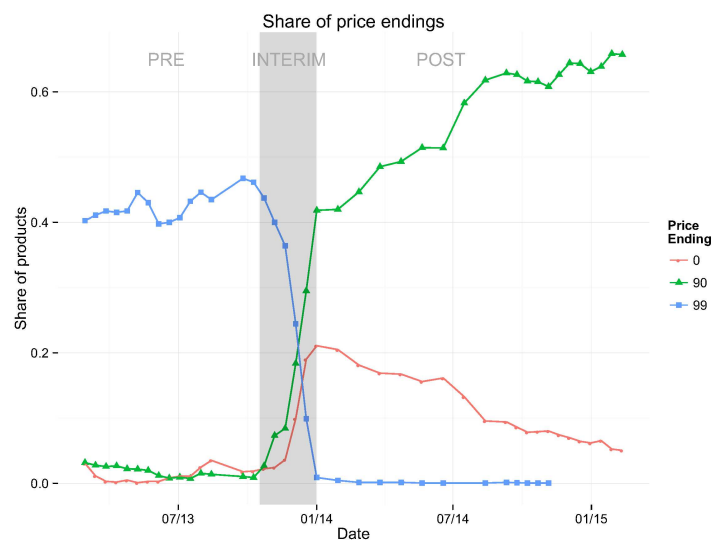


Figure 2.1: Minimal levels of left-digit bias $\theta_0(\epsilon)$ above which 00-ending is suboptimal

The figure shows the minimal levels of left-digit bias θ above which 00-ending prices are never optimal. The values are a function of the price elasticity and nominal price level. The dark lines are the values in the pre-reform case (where 99-ending on 01-ending prices were allowed), and the light lines are the post-reform values (when 90-ending is the highest price ending and 10-ending is the lowest price ending above 00). To illustrate the effects of the price level, the solid lines are for 2.xx prices and the dashed lines for 5.xx prices.

(a) Shares of products that end with 99, 90 and 00, ICC data



(b) Shares of products that end with 99, 90 and 00, CBS data

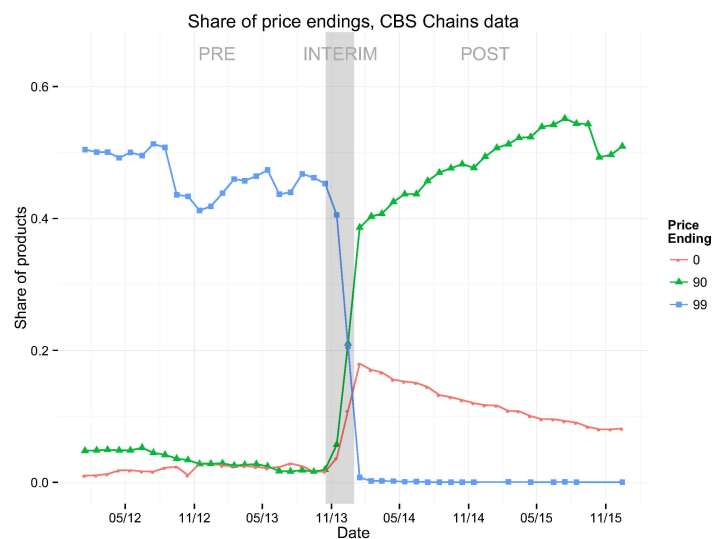


Figure 2.2: Shares of price endings over time

The figures above show that shares of 99-, 90-, and 00-ending prices across products and stores by sampling period. The shaded area represents the time between announcement of the reform on October 17th 2013, and its enactment on January 1st 2014. Squares represent 99-ending prices, triangles 90-ending prices and circles are 00-ending prices. The top panel shows these shares in the ICC data and the bottom panel in the CBS data.

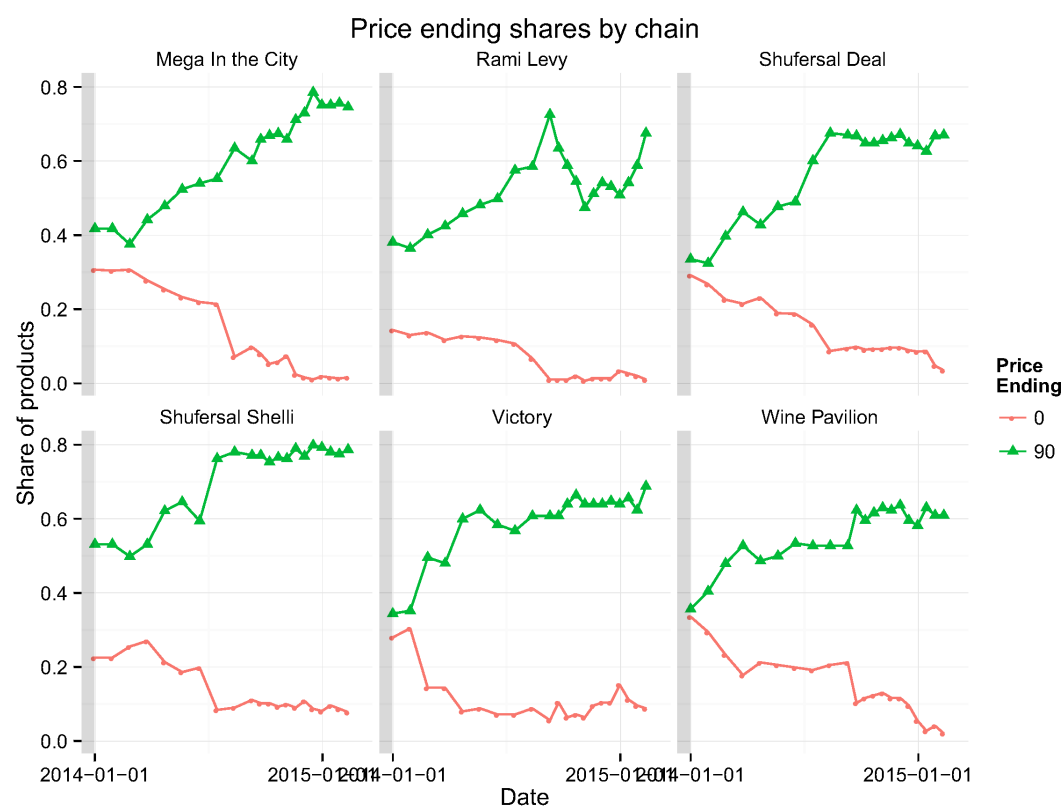
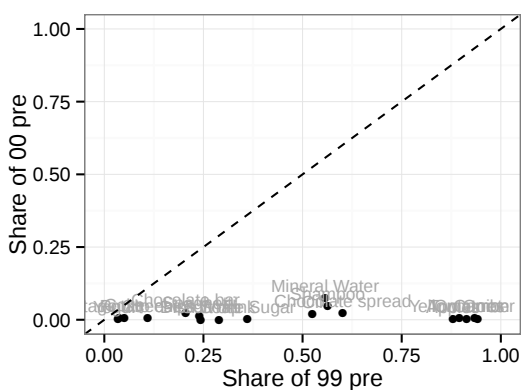


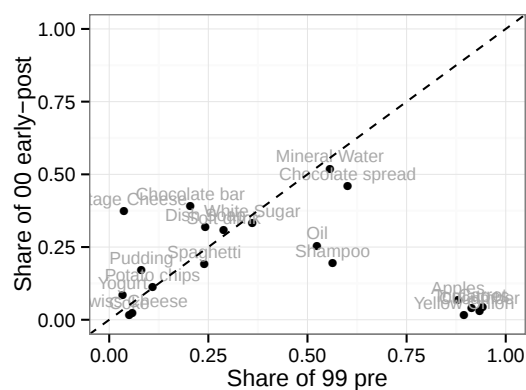
Figure 2.3: Shares of price endings post-reform by chain

The figure shows that shares of 90- and 00-ending prices across products and stores by sampling period only post-reform, broken down by chain. Results are shown for the 6 biggest chains in the data. The shaded area represents the enactment of the reform on January 1st 2014. Triangles represent 90-ending prices and circles are 00-ending prices.

(a) Correlation between share of 99 pre and 00 pre (3-10/2013)



(b) Correlation between share of 99 pre and 00 immediately after policy change (1-2/2014)



(c) Correlation between share of 99 pre and 00 a year after policy change (1-2/2015)

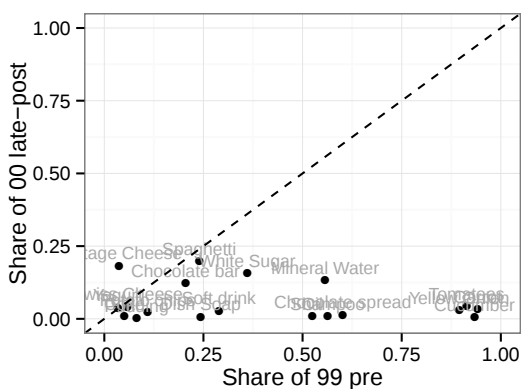


Figure 2.4: Correlation between pre-shares of 99 and post-shares of 90 and 0. Each dot is a product in the ICC data.

The figures show that shares of 00-ending prices at different periods against 99-ending prices at the pre-period, by product. The shares of 00-ending prices per product are calculated before the reform announcement, at the two months following the reform, and at early 2015 a year after the reform.

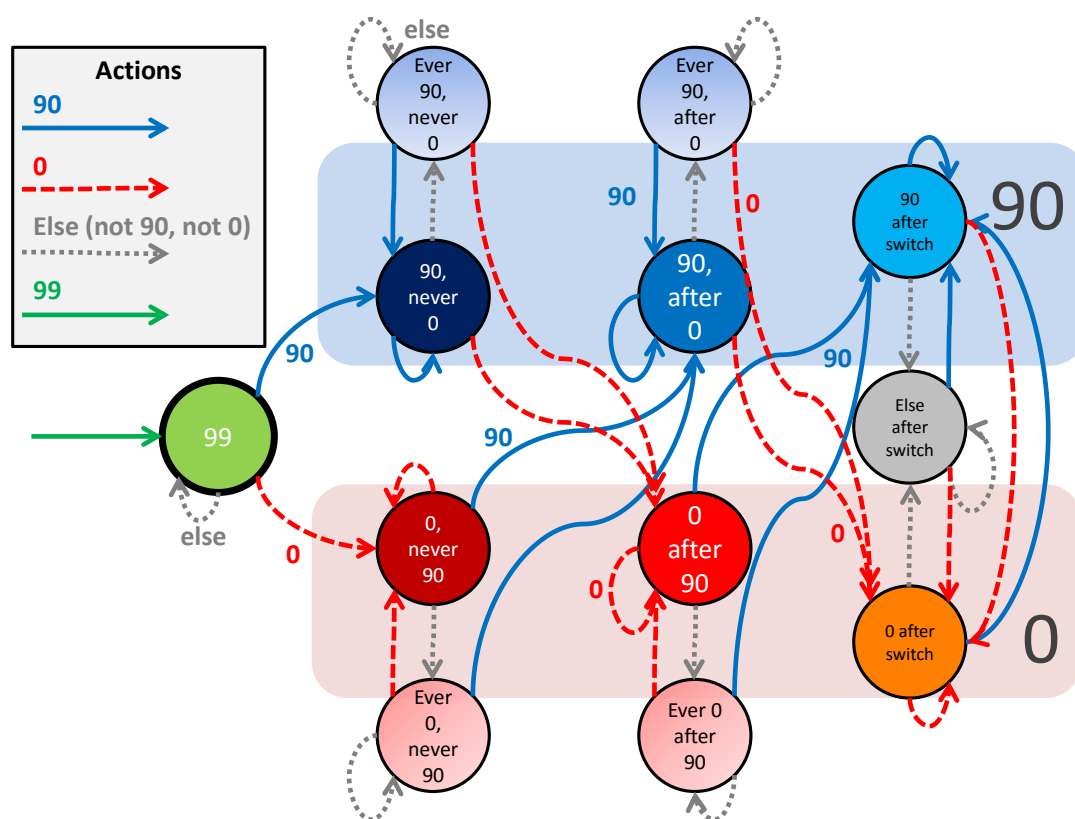


Figure 2.5: Automaton tracking price endings path

The figure illustrates a state-machine that assigns each observation into a “price-ending path” as a function of 4 possible price endings - 99, 90, 00, or other. For example, a price sequence of $\{4.99, 4.90, 5.00, 5.90\}$ will have the corresponding states $\{99, 90 \text{ never } 0, 0 \text{ after } 90, 90 \text{ after switch}\}$.

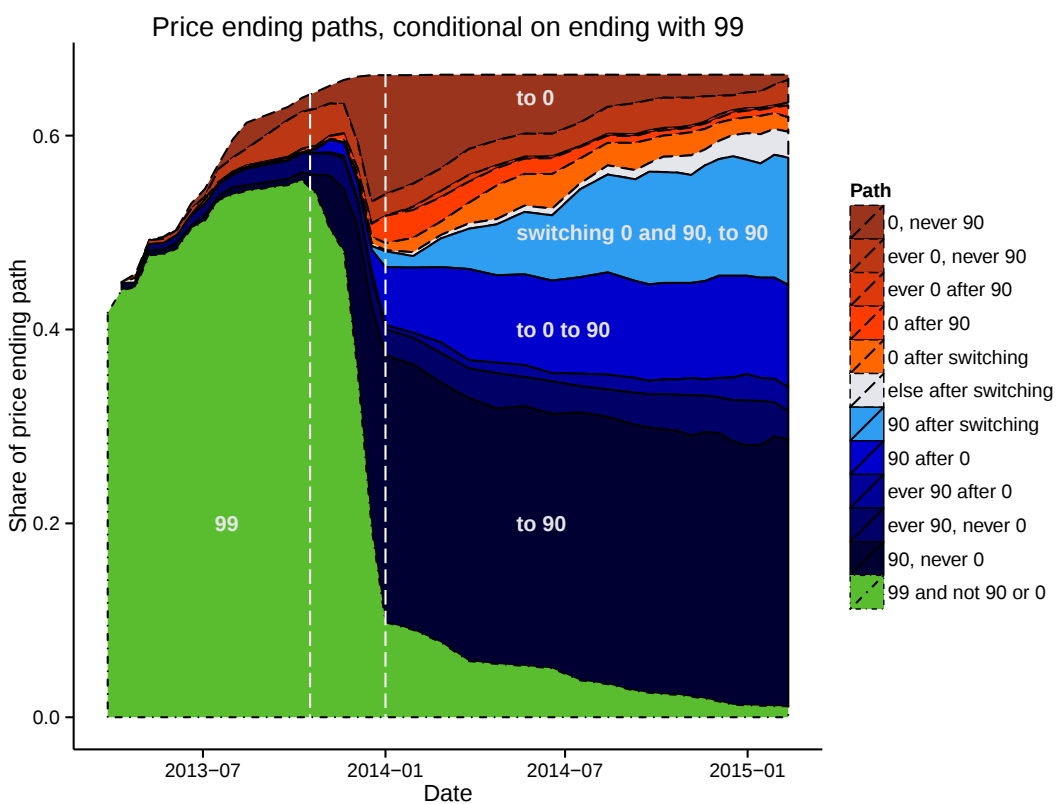
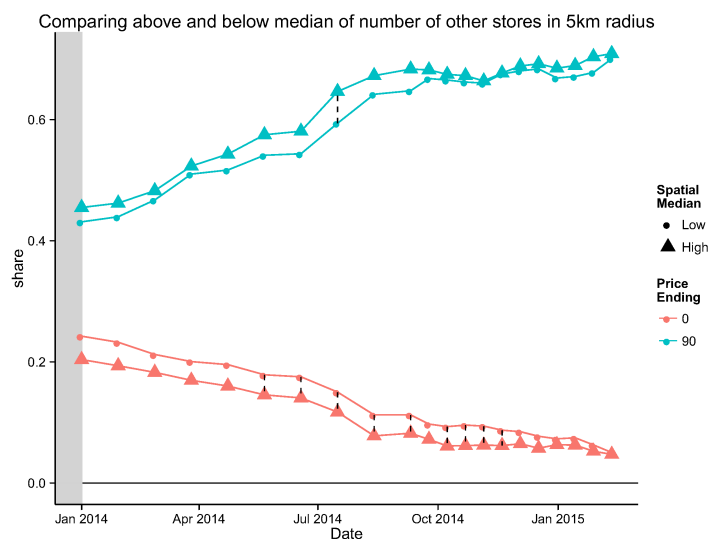


Figure 2.6: Price ending paths of products who ended with 99 before the policy change

The figure shows the shares of each price-ending path (which are described in Figure 2.5). The green area is the 99 or other price endings, the blue shades are the 90-ending prices, and the red shades are the 00-ending prices.

(a) Shares of 00 and 90 for stores above and below median of spatial competition



(b) Shares of 00 and 90 for “Shufersal Deal” stores above and below median of spatial competition

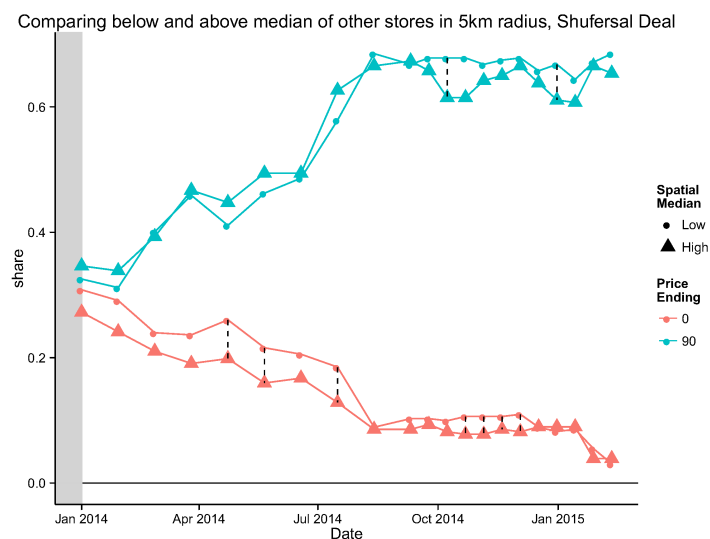


Figure 2.7: Spatial competition and shares of price ending

The figure shows the shares of 00- and 90-ending prices at the post-period for stores that are above and below the median of number of local competitors. The top panel show all stores, and the bottom panel only the stores of the Shufersal Deal chain.

Table 2.1: Summary Statistics of Prices Data

Variable	ICC Raw	ICC Post	ICC Main	CBS Raw	CBS Main	SN Raw	SN Main
No. of Products	67	25	21	233	171	6	6
No. of Stores	659	121	109	155	99	295	278
No. of Chains	38	19	17	—	—	12	12
Sampling Periods	41	21	41	48	48	1095	939
First Observation	Feb 2013	Jan 2014	Feb 2013	Jan 2012	Jan 2012	1/1/2013	1/1/2013
Last Observation	Feb 2015	Feb 2015	Feb 2015	Dec 2015	Dec 2015	12/31/2015	12/31/2015
Daily Units Sold	—	—	—	—	—	18.66	18.94
Mean Price	12.99	7.9	6.88	10.48	10.04	9.01	9.35
Price SD	1.6	0.96	0.88	1.88	1.3	2.8	2.69
Pre 99 Share	0.38		0.43	0.36	0.47	0.19	0.30
Post 90 Share	0.52	0.62	0.58	0.39	0.48	0.33	0.43
Post 00 Share	0.12	0.11	0.11	0.09	0.12	0.12	0.14
Observations	472,572	53,886	82,820	292,471	156,336	1,738,380	774,507

The table shows summary statistics of the three datasets used in the paper, in their raw and cleaned form. ICC, Israeli Consumer Council, collected a panel of prices at the item-store-biweek level. CBS, Israel's Central Bureau of Statistics, collected prices for the consumer price index at the product type-store-month level. SN, StoreNext, is a market research firm, collecting scanner data from retailers including quantities and revenues at the store-day-item level.

	Pre-reform 99-endings share			
	(1)	(2)	(3)	(4)
Pre-reform 00-endings share [1/2013 - 10/2013]	-0.005** (0.002)	-0.002 (0.003)	-0.003 (0.003)	-0.012** (0.005)
Early-post 00-endings share [1/2014 - 2/2014]	-0.139*** (0.027)	-0.134*** (0.027)	-0.144*** (0.028)	0.007 (0.044)
Late-post 00-endings share [1/2015 - 2/2015]	-0.055*** (0.012)	-0.053*** (0.012)	-0.055*** (0.013)	0.001 (0.024)
Chain FE		X		
Store FE			X	
Product FE				X
N	1846	1846	1846	1846

Table 2.2: Correlation between 99-endings pre-reform and 00-endings prices in different periods

Each cell in the table is the coefficient on pre-reform 99-endings share (the shares of 99-endings prices for a product p in a store s before the reform was announced, 1-10/2013), where the dependent variable is on the leftmost column. An observation is a product in a store at a specific period. For example, the bottom cell of column (3) is the coefficient β from the regression $Late00_{ps} = \beta \cdot Pre99_{ps} + Store_{ps} + \epsilon_{ps}$. Standard errors are clustered at the store level. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$

	(1)	(2)	(3)	(4)	(5)	(6)
α^{00}	0.018 (0.014)	0.044 (0.019)	0.218 (0.020)	-0.015 (0.020)	0.039 (0.021)	0.230 (0.028)
α^{90}	0.006 (0.019)	0.055 (0.013)	0.203 (0.020)	0.026 (0.018)	0.057 (0.018)	0.196 (0.025)
α^{99}	-0.043 (0.014)	-0.023 (0.014)	-0.199 (0.016)	-0.042 (0.018)	-0.028 (0.018)	-0.196 (0.022)
$p(\alpha^{90} = \alpha^{00})$	0.541	0.570	0.504	0.077	0.507	0.268
$p(\alpha^{99} = \alpha^{00})$	0.040	0.005	0.000	0.324	0.019	0.000
sample	5+ spells	5+	small change 5+	15+ spells	15+	small change 15+
IV		X	X		X	X
N	394,471	394,471	42,101	291,481	291,481	30,836

Table 2.3: Price-ending effects

The table shows the results from 2.4 under different specifications and samples. Columns (1)-(3) include all instances in which the price is the same for at least 5 consecutive days (columns (4)-(6) for at least 15 days) to circumvent measurement issues. “small change” in columns (3) and (6) further restricts the sample to include instances in which most of the 99-ending prices pre-reform were close to the 00-ending price post-reform. “IV” means that for each product the log(price) is instrumented by the log of the average price of the same product within the same day but in other stores of the same chain. If “IV” is not used, as in columns (1) and (4) then it is OLS of the log(price) interacted with a product in a store.

Bibliography

- Aalto-Setälä, V. (2005). How do markets behave? the adjustment of price endings. *Journal of Product & Brand Management*, 14(7), 455–459.
- Agarwal, S., Chomsisengphet, S., Mahoney, N., & Stroebe, J. (2015). Regulating consumer financial products: Evidence from credit cards. *Quarterly Journal of Economics*, 130(1), 111–164.
- Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools: The economics of manipulation and deception*. Princeton University Press.
- Anderson, E. T., & Simester, D. I. (2003). Effects of \$9 price endings on retail sales: Evidence from field experiments. *Quantitative Marketing and Economics*, 1(1), 93–110.
- Anderson, E., Jaimovich, N., & Simester, D. (2015). Price stickiness: Empirical evidence of the menu cost channel. *Review of Economics and Statistics*, 97(4), 813–826.
- Ashton, L. (2014). Left-digit bias and inattention in retail purchases: Evidence from a field experiment. Available at SSRN: <https://ssrn.com/abstract=2538816>.
- Ater, I., & Gerlitz, O. (2017). Round prices and price rigidity: Evidence from outlawing odd prices. *Journal of Economic Behavior & Organization*, 144, 188–203.
- Azar, O. H. (2007). Relative thinking theory. *The Journal of Socio-Economics*, 36(1), 1–14.
- Basu, K. (1997). Why are so many goods priced to end in nine? and why this practice hurts the producers. *Economics Letters*, 54(1), 41–44.
- Basu, K. (2006). Consumer cognition and pricing in the nines in oligopolistic markets. *Journal of Economics & Management Strategy*, 15(1), 125–141.
- Bizer, G. Y., & Schindler, R. M. (2005). Direct evidence of ending-digit drop-off in price information processing. *Psychology and Marketing*, 22(10), 771–783.
- Blattberg, R. C., & Neslin, S. A. (1990). *Sales promotion: Concepts, methods, and strategies*. Prentice Hall.
- Blattberg, R. C., & Wisniewski, K. J. (1988). *Issues in modeling store level scanner data*. University of Chicago, Center for Research in Marketing.
- Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4), 1351–1408.
- Bray, J. P., & Harris, C. (2006). The effect of 9-ending prices on retail sales: A quantitative uk based field study. *Journal of Marketing Management*, 22(5-6), 601–617.

- Bronnenberg, B. J., Dubé, J.-P., Gentzkow, M., & Shapiro, J. M. (2015). Do pharmacists buy bayer? informed shoppers and the brand premium. *The Quarterly Journal of Economics*, 130(4), 1669–1726.
- Bushong, B., Rabin, M., & Schwartzstein, J. (2015). A model of relative thinking. *Unpublished manuscript, Harvard University, Cambridge, MA*.
- Busse, M. R., Lacetera, N., Pope, D. G., Silva-Risso, J., & Sydnor, J. R. (2013). Estimating the effect of salience in wholesale and retail car markets. *American Economic Review*, 103(3), 575–79.
- Carver, J. R., & Padgett, D. T. (2012). Product category pricing and future price attractiveness: 99-ending pricing in a memory-based context. *Journal of Retailing*, 88(4), 497–511.
- Cavallo, A. (2016). Scraped data and sticky prices. *Review of Economics and Statistics*, (0).
- Cavallo, A., & Rigobon, R. (2016). The billion prices project: Using online prices for measurement and research. *Journal of Economic Perspectives*, 30(2), 151–78.
- Chava, S., & Yao, V. W. (2017). Cognitive reference points, the left-digit effect, and clustering in housing markets.
- Chen, Y., Iyer, G., & Pazgal, A. (2010). Limited memory, categorization, and competition. *Marketing Science*, 29(4), 650–670.
- Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4), 1145–1177.
- Cho, S., & Rust, J. (2008). Is econometrics useful for private policy making? a case study of replacement policy at an auto rental company. *Journal of econometrics*, 145(1-2), 243–257.
- Cho, S., & Rust, J. (2010). The flat rental puzzle. *The Review of Economic Studies*, 77(2), 560–594.
- Conlon, C. T., & Rao, N. (2016). Discrete prices and the incidence and efficiency of excise taxes.
- Dalrymple, D. J., & Haines Jr, G. H. (1970). A study of the predictive ability of market period demand-supply relations for a firm selling fashion products. *Applied Economics*, 1(4), 277–285.
- DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2), 315–372.
- DellaVigna, S. (2018). Chapter 7 - structural behavioral economics. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of behavioral economics - foundations and applications 1* (Vol. 1, pp. 613–723). Handbook of Behavioral Economics: Applications and Foundations 1. doi:<https://doi.org/10.1016/bs.hesbe.2018.07.005>
- DellaVigna, S., & Gentzkow, M. (2017). Uniform pricing in us retail chains.
- DellaVigna, S., & Malmendier, U. (2004). Contract design and self-control: Theory and evidence. *Quarterly Journal of Economics*, 119(2), 353–402.
- DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *American Economic Review*, 96(3), 694–719.

- Dube, A., Manning, A., & Naidu, S. (2017). Monopsony and employer mis-optimization account for round number bunching in the wage distribution. *Unpublished manuscript*.
- Eichenbaum, M., Jaimovich, N., & Rebelo, S. (2011). Reference prices, costs, and nominal rigidities. *The American Economic Review*, 101(1), 234–262.
- Eichenbaum, M., Jaimovich, N., Rebelo, S., & Smith, J. (2014). How frequent are small price changes? *American Economic Journal: Macroeconomics*, 6(2), 137–155.
- Einav, L., Leibtag, E., & Nevo, A. (2010). Recording discrepancies in nielsen homescan data: Are they present and do they matter? *QME*, 8(2), 207–239.
- El Sehity, T., Hoelzl, E., & Kirchler, E. (2005). Price developments after a nominal shock: Benford's law and psychological pricing after the euro introduction. *International Journal of Research in Marketing*, 22(4), 471–480.
- Farhi, E., & Gabaix, X. (2015). *Optimal taxation with behavioral agents*. National Bureau of Economic Research.
- Foxall, G. R., Goldsmith, R. E., & Brown, S. (1998). *Consumer psychology for marketing*. Cengage Learning EMEA.
- Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4), 1661–1710.
- Gabaix, X. (2019). Chapter 4 - behavioral inattention. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of behavioral economics - foundations and applications 2* (Vol. 2, pp. 261–343). Handbook of Behavioral Economics: Applications and Foundations 1. doi:<https://doi.org/10.1016/bs.hesbe.2018.11.001>
- Gabaix, X., & Laibson, D. (2009). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics*, 124(3), 969–1010.
- Gedenk, K., & Sattler, H. (1999). The impact of price thresholds on profit contribution should retailers set 9-ending prices? *Journal of Retailing*, 75(1), 33–57.
- Ginzberg, E. (1936). Customary prices. *The American Economic Review*, 26(2), 296–296.
- Goldfarb, A., & Xiao, M. (2011). Who thinks about the competition? managerial ability and strategic entry in us local telephone markets. *American Economic Review*, 101(7), 3130–61.
- Gopinath, G., Gourinchas, P.-O., Hsieh, C.-T., & Li, N. (2011). International prices, costs, and markup differences. *The American Economic Review*, 101(6), 2450–2486.
- Grubb, M. D., & Osborne, M. (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review*, 105(1), 234–271.
- Hackl, F., Kummer, M. E., & Winter-Ebmer, R. (2014). 99 cent: Price points in e-commerce. *Information Economics and Policy*, 26, 12–27.
- Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3), 1311–1353.
- Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods* (pp. 207–248). University of Chicago Press.

- Hausman, J., Leonard, G., & Zona, J. D. (1994). Competitive analysis with differentiated products. *Annales d'Economie et de Statistique*, 159–180.
- Heidhues, P., & Kőszegi, B. (2018). Chapter 6 - behavioral industrial organization. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of behavioral economics - foundations and applications 1* (Vol. 1, pp. 517–612). Handbook of Behavioral Economics: Applications and Foundations 1. doi:<https://doi.org/10.1016/bs.hesbe.2018.07.006>
- Hendel, I., Lach, S., & Spiegel, Y. (2017). Consumers' activism: The cottage cheese boycott. *The RAND Journal of Economics*, 48(4), 972–1003.
- Hilger, N. (2018). Heuristic thinking in the market for online subscriptions. *Available at SSRN 3296698*.
- Hirshman, S., Pope, D., & Song, J. (forthcoming). Mental budgeting vs. relative thinking. *American Economic Association Papers and Proceedings*.
- Hossain, T., & Morgan, J. (2006). ...plus shipping and handling: Revenue (non) equivalence in field experiments on ebay. *The B.E. Journal of Economic Analysis & Policy*, 5(2), 1–30.
- Jiang, Z. (2018). An empirical bargaining model with digit bias - a study on auto loan monthly payments. *Unpublished manuscript*.
- Kalyanam, K., & Shively, T. S. (1998). Estimating irregular pricing effects: A stochastic spline regression approach. *Journal of Marketing Research*, 16–29.
- Kleven, H. J., & Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2), 669–723.
- Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, 102(5), 2206–2236.
- Levy, D., Lee, D., Chen, H., Kaufmann, R. J., & Bergen, M. (2011). Price points and price rigidity. *Review of Economics and Statistics*, 94(4), 1417–1431.
- Macé, S. (2012). The impact and determinants of nine-ending pricing in grocery retailing. *Journal of Retailing*, 88(1), 115–130.
- Matějka, F. (2015). Rationally inattentive seller: Sales and discrete pricing. *The Review of Economic Studies*, 83(3), 1125–1155.
- Montgomery, A. L. (1997). Creating micro-marketing pricing strategies using supermarket scanner data. *Marketing science*, 16(4), 315–337.
- Nakamura, E., & Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4), 1415–1464.
- Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2), 307–342.
- Nevo, A. (2011). Empirical models of consumer behavior. *Annual Review of Economics*, 3(1), 51–75.
- Repetto, L., & Solis, A. (2018). The price of inattention: Evidence from the swedish housing market.

- Rossi, P. E. (2014). Even the rich can make themselves poor: A critical examination of iv methods in marketing applications. *Marketing Science*, 33(5), 655–672.
- Rossi, P. E., Allenby, G. M., & McCulloch, R. (2012). *Bayesian statistics and marketing*. John Wiley & Sons.
- Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3), 180–212.
- Schindler, R. M., & Chandrashekar, R. (2004). Influence of price endings on price recall: A by-digit analysis. *Journal of Product & Brand Management*, 13(7), 514–524.
- Schindler, R. M., & Kibarian, T. M. (1996). Increased consumer sales response though use of 99-ending prices. *Journal of Retailing*, 72(2), 187–199.
- Schindler, R. M., & Kirby, P. N. (1997). Patterns of rightmost digits used in advertised prices: Implications for nine-ending effects. *Journal of Consumer Research*, 24(2), 192–201.
- Schindler, R. M., & Wiman, A. R. (1989). Effects of odd pricing on price recall. *Journal of Business Research*, 19(3), 165–177.
- Snir, A., Levy, D., & Chen, H. A. (2017). End of 9-endings, price recall, and price perceptions. *Economics Letters*, 155, 157–163.
- Spiegler, R. (2011). *Bounded rationality and industrial organization*. OUP USA.
- Stiving, M. (2000). Price-endings when prices signal quality. *Management Science*, 46(12), 1617–1629.
- Stiving, M., & Winer, R. S. (1997). An empirical analysis of price endings with scanner data. *Journal of Consumer Research*, 24(1), 57–67.
- Strulov-Shlain, A. (2019a). Firm sophistication: A natural experiment on price endings. *Unpublished manuscript*.
- Strulov-Shlain, A. (2019b). More than a penny’s worth: Left-digit bias and firm pricing. *Unpublished manuscript*.
- Taubinsky, D., & Rees-Jones, A. (2016). *Attention variation and welfare: Theory and evidence from a tax salience experiment* (Working Paper No. 22545). National Bureau of Economic Research. doi:10.3386/w22545
- Thomas, M., & Morwitz, V. (2005). Penny wise and pound foolish: The left-digit effect in price cognition. *Journal of Consumer Research*, 32(1), 54–64.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.
- Varian, H. R. (1980). A model of sales. *The American Economic Review*, 70(4), 651–659.
- Wansink, B., Kent, R. J., & Hoch, S. J. (1998). An anchoring and adjustment model of purchase quantity decisions. *Journal of Marketing Research*, 71–81.

Appendix A

Appendix to Chapter 1

A.1 Proofs

Proof - Proposition 1.2

To see that equation 1.3 describes the optimal price, we will proceed in the following steps.

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

$$\begin{aligned}\frac{\partial \Pi}{\partial p} = 0 &= A\epsilon((1-\theta)p + \theta p_1)^{\epsilon-1}(1-\theta)(p-c) + A((1-\theta)p + \theta p_1)^\epsilon \\ 0 &= \epsilon(1-\theta)(p-c) + (1-\theta)p + \theta p_1 \\ p &= c \frac{\epsilon}{1+\epsilon} - p_1 \frac{\theta}{1-\theta} \frac{1}{1+\epsilon} \\ p &= \left(c - \frac{\theta}{1-\theta} \frac{q_1+1}{\epsilon} \right) \frac{\epsilon}{1+\epsilon}\end{aligned}$$

Where the last transition is due to that $p_1 = q_1 + 1$ given a range of costs as in proposition 1.3. So, first, for any cost c , we can find a local maximum being the interior solution, as presented in the second row of equation 1.3. Note that this solution is strictly increasing in c , as long as p_1 does not change.

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 $q = p_1 - 0.01$:

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

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

Once we have shown that q is more profitable than p^* , to see that the optimal price in that case is q itself, note that since demand is locally strictly convex between discontinuity points, if the profit at q is higher than at the interior solution it means that the slope of the isoprofit (in the quantity-price dimension) is less steep than the demand curve, meaning that profits at q must be maximal for these costs. This is only true for θ low enough, such that there is not another point where demand is crossing the isoprofit¹.

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,

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

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 $\Pi(p^*, c) = D(q)(p' - c) > D(q)(q - c)$.

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

Proof - Proposition 1.2

Rewriting equation 1.5,

$$\begin{aligned}
P + \frac{(1-\theta)P + \theta P_1}{\epsilon(1-\theta)} - \frac{((1-\theta)P + \theta P_1)^\epsilon P - ((1-\theta)q + \theta q_1)^\epsilon q}{((1-\theta)P + \theta P_1)^\epsilon - ((1-\theta)q + \theta q_1)^\epsilon} &= 0 \\
P \left(1 + \frac{1}{\epsilon}\right) + \frac{TP_1}{\epsilon} - \frac{(P + TP_1)^\epsilon P - (q + Tq_1)^\epsilon q}{(P + TP_1)^\epsilon - (q + Tq_1)^\epsilon} &= 0 \\
P \left(1 + \frac{1}{\epsilon}\right) + \frac{TP_1}{\epsilon} - \frac{\Lambda^\epsilon P - q}{\Lambda^\epsilon - 1} &= 0 \\
P(1 + \epsilon)(\Lambda^\epsilon - 1) + TP_1\Lambda^\epsilon - TP_1 &= \Lambda^\epsilon P\epsilon - q\epsilon \\
(P + TP_1)\Lambda^\epsilon &= P + TP_1 + (P - q)\epsilon \\
\Lambda^\epsilon &= 1 + \frac{(P - q)\epsilon}{P + TP_1} \quad (\text{A.1})
\end{aligned}$$

where $T \equiv \frac{\theta}{1-\theta}$, the likelihood ratio of being inattentive and $\Lambda \equiv \frac{P+TP_1}{q+Tq_1}$.

The effect of elasticity on nl-price

First, note that $\frac{\partial \Lambda}{\partial \epsilon} = \frac{\partial P}{\partial \epsilon} \frac{1}{q+Tq_1}$. Then, differentiating equation A.1 with respect to ϵ ,

$$\begin{aligned}
(P + TP_1)\Lambda^\epsilon \left(\frac{\epsilon \frac{\partial P}{\partial \epsilon}}{P + TP_1} + \ln \Lambda \right) + \frac{\partial P}{\partial \epsilon} \Lambda^\epsilon &= \frac{\partial P}{\partial \epsilon} + \frac{\partial P}{\partial \epsilon} \epsilon + P - q \\
\frac{\partial P}{\partial \epsilon} (\Lambda^\epsilon - 1)(1 + \epsilon) + (P + TP_1)\Lambda^\epsilon \ln \Lambda &= P - q \\
\frac{\partial P}{\partial \epsilon} &= \frac{P - q - (P + TP_1)\Lambda^\epsilon \ln \Lambda}{(\Lambda^\epsilon - 1)(1 + \epsilon)} > 0
\end{aligned}$$

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

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

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

$$\begin{aligned}
(q + Tq_1)^\epsilon - (\epsilon + 1)(P + TP_1)^\epsilon \ln \left(\frac{P + TP_1}{q + Tq_1} \right) - (P + TP_1)^\epsilon &> 0 \\
\Leftrightarrow 1 > \Lambda^\epsilon (1 + (1 + \epsilon) \ln \Lambda)
\end{aligned}$$

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

Proof of corollary, effect of dollar digit on nl-price

We know that the NLP satisfies:

$$((1 - \theta)P + \theta P_1)^\epsilon \left(-\frac{(1 - \theta)P + \theta P_1}{\epsilon(1 - \theta)} \right) - \quad (\text{A.2})$$

$$((1 - \theta)q + \theta q_1)^\epsilon \left(q - P \cdot \frac{1 + \epsilon}{\epsilon} - \frac{\theta P_1}{(1 - \theta)\epsilon} \right) = 0 \quad (\text{A.3})$$

$$\hat{p}^\epsilon \frac{\hat{p}}{\epsilon(1 - \theta)} + \hat{q}^\epsilon \left(q - p - \frac{\hat{p}}{\epsilon(1 - \theta)} \right) = 0 \quad (\text{A.4})$$

$$\hat{q}^\epsilon (q - p) \epsilon (1 - \theta) = \hat{p} (\hat{q}^\epsilon - \hat{p}^\epsilon) \quad (\text{A.5})$$

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

$$\begin{aligned} (\hat{q} + 1)^\epsilon [(q - p) \epsilon (1 - \theta)] &\lesseqgtr (\hat{p} + 1) ((\hat{q} + 1)^\epsilon - (\hat{p} + 1)^\epsilon) \\ (\hat{q} + 1)^\epsilon [\hat{p} (\hat{q}^\epsilon - \hat{p}^\epsilon) \hat{q}^{-\epsilon}] &\lesseqgtr (\hat{p} + 1) ((\hat{q} + 1)^\epsilon - (\hat{p} + 1)^\epsilon) \\ \frac{\hat{p}}{\hat{p} + 1} &\leq \frac{1 - \left(\frac{\hat{p}+1}{\hat{q}+1}\right)^\epsilon}{1 - \left(\frac{\hat{p}}{\hat{q}}\right)^\epsilon} \end{aligned}$$

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

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

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

The effect of inattention on nl-price

How does higher inattention affects the next lowest price? Differentiate equation A.1 with respect to θ :

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

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

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

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

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

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

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

then it is left to show that

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

or

$$\begin{aligned} q + \Delta + T(1 + q_1) &< (1 + q_1)(q + Tq_1) \\ \Delta + T &< q_1(q + Tq_1) \end{aligned}$$

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.

A.2 Price Competition

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

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

Proposition 1. *There is no pure strategy equilibrium in this game*

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

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

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

An immediate corollary is that in any equilibrium, firms will not price at the point of discontinuity. Taking this simplistic, continuous domain model to our setting, the price of

“1” represents the highest price before the drop in demand, i.e. prices that end with 99 before the policy or 90 post policy. So, that model predicts no pricing at 00.

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

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

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 left².

Also, within each region there are no gaps. Assume there is, i.e. $f(p) = 0$ for all $p \in (\underline{p}, \bar{p})$ where either $\bar{p} < 1$ or $\underline{p} > 1$. Then, it is profitable to shift a probability from $f(\underline{p}) > 0$ 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,

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

²If there is a mass point at 1, then it is profitable to shift it to just below 1 and gain $\theta \frac{1-(1-\phi)^N - (1-\phi)^{N-1}\phi}{N\phi} = \theta \frac{1-(1-\phi)^{N-1}}{N\phi} > 0$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A.3 Multi-product Monopolist

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

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

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

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

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

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

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

If the cross-elasticity is 0, we are back to the single item monopolist, since the problem is separable. Deriving a closed-form solution is possible, but it is not very insightful, in the sense that I do not have any comparative statics to provide. Instead, I solve the problem numerically, and focus attention on the effects of adding cross elasticities on the region of missing prices (numerical, local, comparative statics). Figure A.11 provides an example for such effects. Here, the black curves describe the price versus cost of product i , when product j has a cost of 1.5, and the orange curves are the price of product j . Both items have same price elasticity of -2, and cross-elasticity is either zero (the solid lines) or 0.1 (the dashed line). Margins of product i are higher due to the substitution with product j . Also, the arrows point to the lowest-prices of 4.67 in the single item case, and 4.73 with the cross-elasticity.

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

A.4 Nielsen Data

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

Nielsen Retail Scanner (RMS) and Consumer Panel (HMS) data are provided by the Kilts Center at the University of Chicago. RMS records weekly UPC-store level quantity and revenues. Summary statistics are described in column (1) of Appendix Table 1.1. The sample consists of 28 million observations of 12 products from DellaVigna and Gentzkow, 2017, selling in 11000 stores, with a total of 188 million units sold for \$881 million. The average price is at \$5.51, with a standard deviation of 2.81.

The data is similar to the retailer’s data, with the key advantage being that it includes the clientele and pricing behavior of more than 60 retailers rather than one. Another pragmatic advantage is the longer panel of 9 years instead of 3.5. However, the data have three main disadvantages, and as such is not part of the main analysis. First, prices are rounded such that they are always “to the cent”, thus not allowing to infer if an observation represents a fictitious price. Second, Nielsen measurement weeks start on Sunday, and are not always aligned with a retailer’s price changing frequency. For example, some chains change their prices weekly but do so on Wednesdays, so if a price changed the Nielsen price will mix sales with one price for Sunday-Tuesday and another for Wednesday-Saturday. Measurement misalignment leads to far more fictitious prices than in the retailer data. Finally, only the weekly average price is observed, excluding a simple way to infer if a price is on-sale or not. Additional variables, “promotion” and “display”, are available but are missing for 80% of the observations.

To solve the issues of fictitious prices and sale prices I use a new, yet simple, data cleaning technique. If an observed price of a unique product in a certain store is the result of averaging different prices, it is unlikely to be identical to the price in the following week ³. For example, imagine there is a discount to club members only, so some consumers pay \$4.99 and others \$4.29. In order for the “weekly” price to be the same across weeks, the share of club members among weekly consumers must be kept exactly the same. Therefore, I define for each product-store the spell length of each price and exclude 1-week spells. Longer spells are also more likely to be regular prices, or conversely, shorter spells are more likely to be on-sale prices. Therefore, while I can not directly observe if a product is on-sale in a certain week, I proxy for it with the spell length. The spell length proxy has the advantage of not inferring an on-sale by price, which is problematic since regular prices also change.

³Mostly true for a popular product purchased in multiple units each week, otherwise might be likely.

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

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

A.5 Bias Heterogeneity

Armed with estimates of the bias at the product level, we can ask how does the bias change according to the product characteristics? A few mechanisms are interesting to look at. First, a mechanism of relative thinking might be in place. The idea was introduced by Tversky and Kahneman, 1981 with the famous calculator-jacket example. Namely, 68% of respondents take the same hypothetical effort for a \$5 discount if it is out of a \$15 item, while only 29% if it is out of a \$125 item. Needless to say, this kind of behavior is not consistent with utility maximization. While different theories were developed (e.g., Azar, 2007; Bushong, Rabin, and Schwartzstein, 2015), field evidence for the effect is rare (with the recent exception of Hirshman, Pope, and Song, forthcoming). In this left-digit bias setting, a manifestation of relative thinking is if the bias is higher for high ticket items, everything else constant. Meaning, if a cent out of a \$2 item weighs more heavily than a cent out of an \$8 item. Already hinted in Figure 1.4, 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 trip⁴. 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, Dubé, Gentzkow, and Shapiro, 2015).

Since I have various measures of the bias from the three methods discussed above, each with its benefits and drawbacks, I use a simple mean as the bias measure. This is a crude way of minimizing sampling error in the estimates. The simple correlates appear in Figure A.9. As hinted above, price (panel a), is positively correlated with the bias; demand is negatively correlated (panel b); and education and income are less clearly correlated (panels c and d).

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

A.6 Welfare

How does the bias affects welfare? Section 1.6 discusses the basic concepts, of the demand and price effects. Here I will present a more formal treatment, followed by simulations demonstrating the comparative statics.

⁴I 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.

Constant elasticity

For exposition purposes, I begin with the monopolistic constant elasticity demand benchmark of the paper. Recall that demand is a function of price and left-digit bias, $D(p, \theta) = A\hat{p}^\epsilon$.

- Profits given a price p_* and unit costs c are

$$\Pi = A\hat{p}_*^\epsilon (p_* - c)$$

since $p_* = (1 - \theta)p_* + \theta(\lfloor p_* \rfloor + \Delta)$, if $\Delta = 0$ then the price with bias is always perceived as lower than it is $p_*(\theta) \leq p_*(0)$, and in contrast if $\Delta = 0.99$ then $p_*(\theta) \geq p_*(0)$. So, for $\Delta = 0$ ($\Delta = 0.99$) demand is higher (lower) for every price, and therefore profits are higher (lower) by assumption.

- Consumer surplus is the value created by the transaction. Meaning, the area above the price, p_* , and potentially demanded quality if the price were perceived correctly. That is,

$$CS = \int_{D(\infty, \theta)}^{D(p_*, \theta)} (D^{-1}(q, \theta = 0) - p_*) dq = \int_0^{A\hat{p}_*^\epsilon} \left(\left(\frac{q}{A} \right)^{\frac{1}{\epsilon}} - p_* \right) dq$$

which can be broken down to the surplus absent a bias and the transfer from consumers to producers from overconsumption if there is any

$$= \int_0^{A\hat{p}_*^\epsilon} \left(\left(\frac{q}{A} \right)^{\frac{1}{\epsilon}} - p_* \right) dq - \int_{A\hat{p}_*^\epsilon}^{A\hat{p}_*^\epsilon} \left(p_* - \left(\frac{q}{A} \right)^{\frac{1}{\epsilon}} \right) dq$$

which equals in total

$$\begin{aligned} CS &= A \left(\frac{\epsilon}{\epsilon + 1} \hat{p}_*^{\epsilon+1} - p_* \cdot \hat{p}_*^\epsilon \right) \\ &= A \left(\frac{\epsilon}{\epsilon + 1} [(1 - \theta)p_* + \theta(\lfloor p_* \rfloor + \Delta)]^{\epsilon+1} - p_* \cdot [(1 - \theta)p_* + \theta(\lfloor p_* \rfloor + \Delta)]^\epsilon \right) \end{aligned}$$

- Deadweight loss is the amount of surplus that could have been created, i.e. the area between true demand and demand at cost

$$DWL = \int_{D(p_*, \theta)}^{D(c, \theta)} \left(\left(\frac{q}{A} \right)^{\frac{1}{\epsilon}} - c \right) dq$$

If one can see beyond the algebra, the above make clear that the price level p_* is the key component, together with the left-digit bias that sets the distortions in demand (depending on Δ) and transfers from consumers to producers.

p_* , the set price, does not depend on θ , the consumers left-digit bias, but on $\hat{\theta}$, the bias the firm perceives to exist. Such that p_* solves

$$p_* (\hat{\theta}, c, \epsilon) = \arg \max_p \left[(1 - \hat{\theta})p + \hat{\theta}(\lfloor p \rfloor + \Delta) \right]^\epsilon (p - c)$$

As shown in the paper in equation 1.3, the price is shifted up in the interior solutions region and down to a lower 99-ending price otherwise. Meaning that we know that the effect of the bias on prices are mixed in sign

$$p_*(\theta, c, \epsilon) \leq p_*(0, c, \epsilon)$$

Further, if $\hat{\theta} < \theta$ then: If $p_*(\theta, c, \epsilon) > p_*(0, c, \epsilon) \Rightarrow p_*(\theta, c, \epsilon) > p_*(\hat{\theta}, c, \epsilon) > p_*(0, c, \epsilon)$; and if $p_*(\theta, c, \epsilon) < p_*(0, c, \epsilon)$ and is 99-ending, then $p_*(\hat{\theta}, c, \epsilon) \geq p_*(\theta, c, \epsilon)$. Meaning, under bias underestimation of firms, the effects on prices are also mixed in sign. Note that these effects are not affected by Δ .

Together, these mean that the effects on CS and DWL are ambiguous in sign as they depend on the underlying distribution of costs and elasticities.

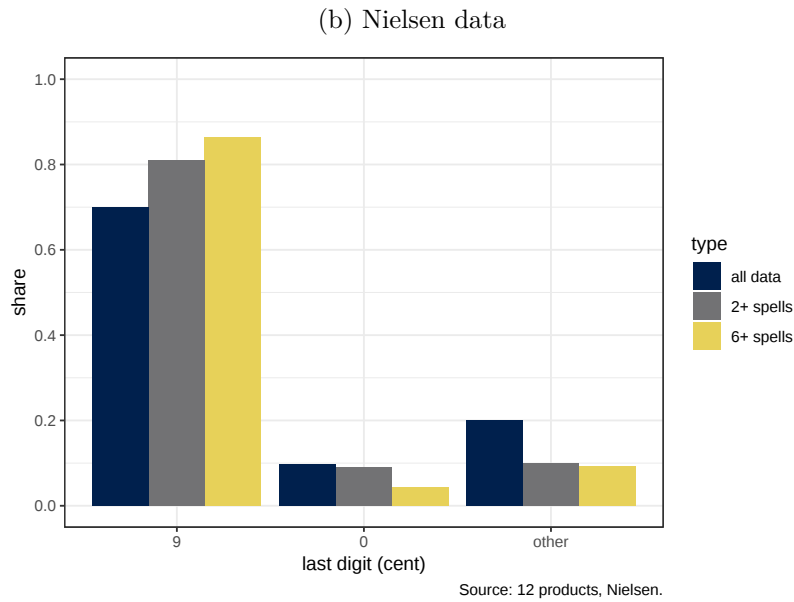
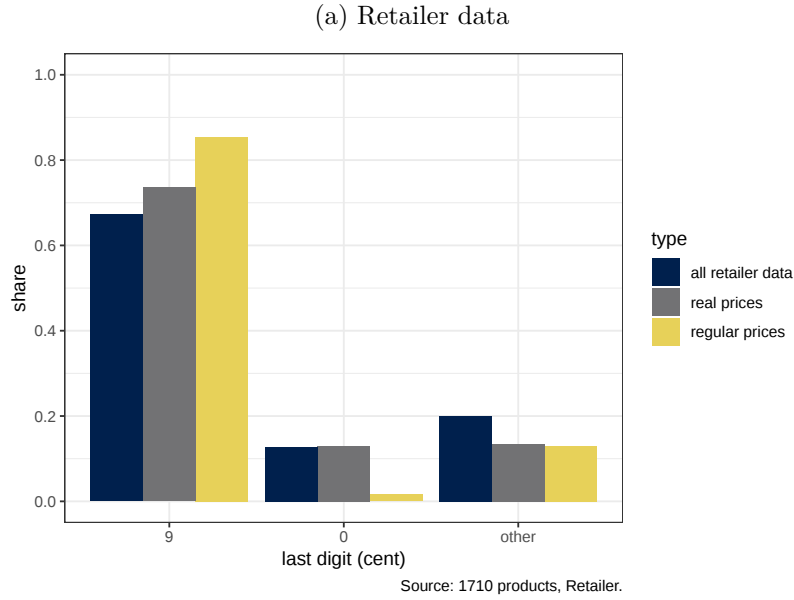
Simulations and other demand structures

The mixed effects are not sensitive to the demand structure in a monopolistic setting. The reason being that the same forces operate for any smooth demand with discontinuities – left-digit bias as modeled here, as a mix of the true price with a focal price ending, will shift demand reducing within-digit sensitivity and creating sharp incentives to deviate to the lower 99-ending price. Profits, as with constant elasticities, will be such that demand is affected by perceived prices but consumers pay the actual price. Consumer Surplus is the area between “true” demand and the price, so the shape of the demand curve will change above the price, but the effects of a higher price p_* remain a reduction in surplus, while a lower p_* means a mix of benefits and costs to consumers. Similarly, DWL is affected by whether or not the price is pushed down due to the bias.

Another question is whether this mixtures matter in practice or just in theory. While true demand and cost structure are not observable, we can simulate different distributions of costs and demand structures to at least see if the different components are indeed changing signs, and how. Table A.4 summarizes the signs of effects in the following exercise: I simulate 10,000 consumers with individual valuation for a product drawn from a distribution F . A specific distribution determines the local elasticity at each price point. Consumers have left-digit bias θ with focal price ending of Δ and the firm has perceived left-digit bias of $\hat{\theta} \leq \theta$. I then simulate a distribution of 500 costs on a uniform distribution, solve for the price and surpluses at each set of parameters $(\theta, \hat{\theta}, c, \Delta)$. I average across costs to obtain a number at each tuple $(\theta, \hat{\theta}, \Delta)$.

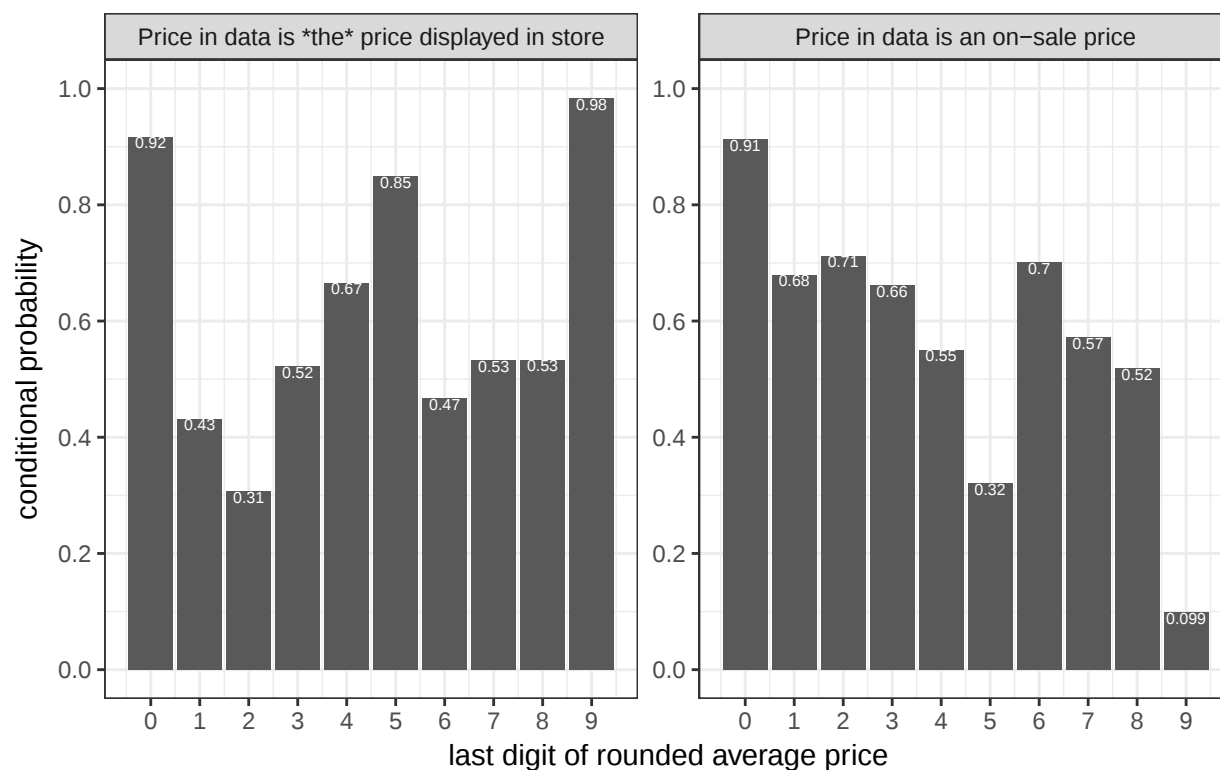
One exercise is looking at the values at $\hat{\theta} = \theta > 0$ compared to a baseline of $\theta = \hat{\theta} = 0$ to see the effects of higher left-digit bias when firms are sophisticated. The second exercise, is to compare the values fixing θ and only changing $\hat{\theta}$ to 0 (meaning the firm prices as if there is no bias at all) and to 0.005 (underestimation of the bias, roughly as estimated in Section 1.5).

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



The figures show the shares of 9-ending prices, 0-ending prices and others in the Retailer data (panel A) and Nielsen data (panel B), and how these change by sample selection. The dark blue bars are the data before cleaning (corresponding to column (2) in Table 1.1 and column (1) in A.1). The middle green bars are the shares of “real” prices in the data (columns (3) and (2)) defined as to-the-cent prices in the Retailer data and at least 2 week price spells in Nielsen. The light yellow bars are the shares for “regular” prices (columns (4) and (3)), defined as instances in which the net price equals the regular price in the Retailer data and price spells of at least 6 weeks in Nielsen.

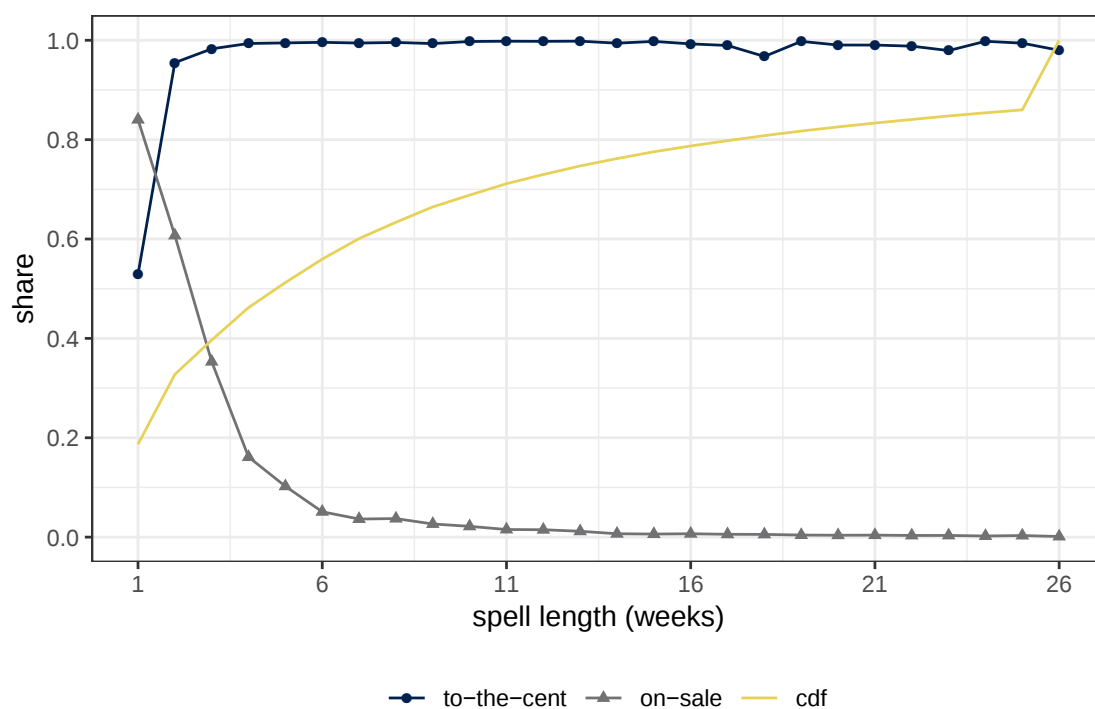
Figure A.2: Data description: “real” and “on-sale” by last digit of observed price



Based on 1710 products, 52M observations. `on-sale` = weekly average price is lower than regular price.
 `real price` = the weekly average price is `to the cent`. Shares of price endings: 9: 67%, 0: 13%, 5: 6%, others <2%.
 source: retailer data.

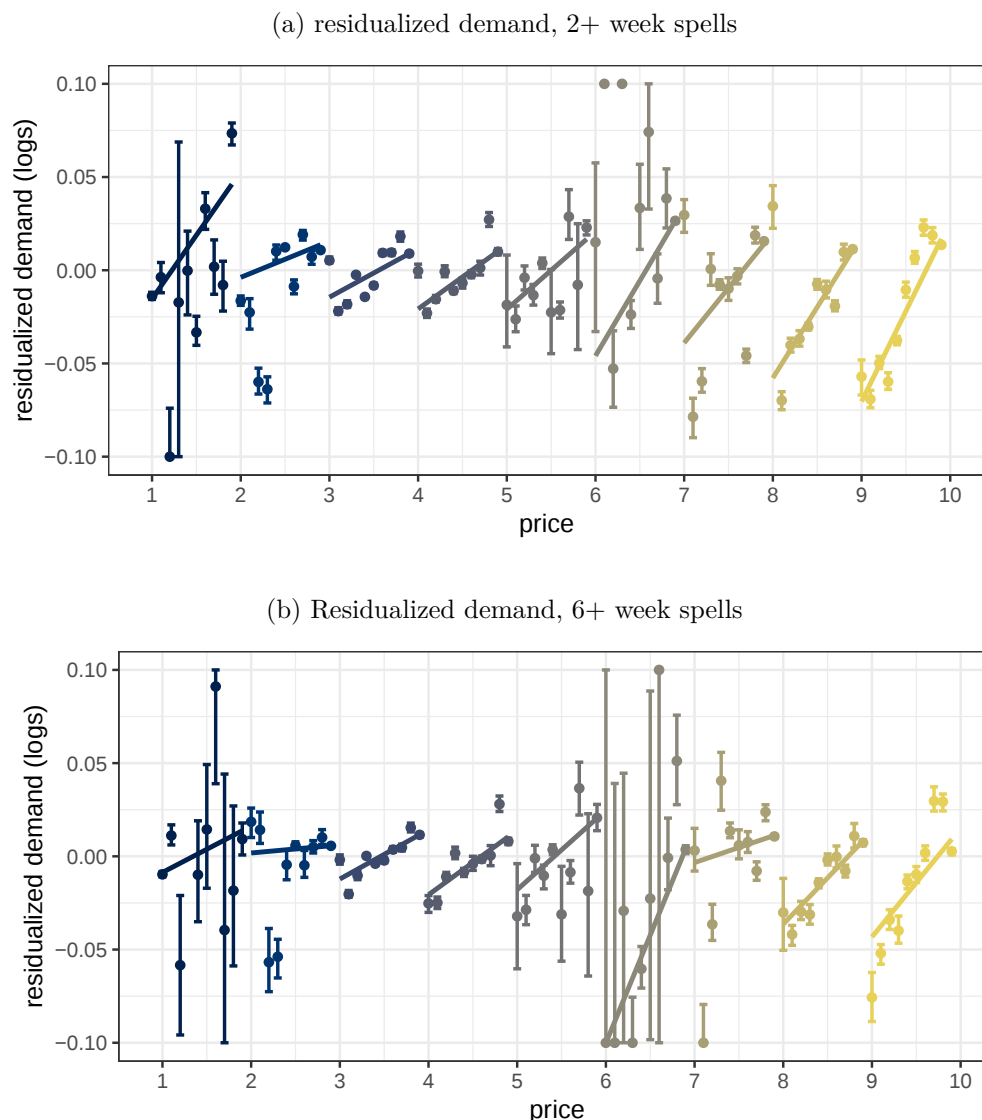
The figure shows the shares of observed prices that are to-the-cent (left) and on-sale (right), conditional on the last digit of the (rounded to the cent) price. For example, on the left, prices ending with 9 are 98% to-the-cent while 2-ending prices are to-the-cent in 31% of observations. Meaning, for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

Figure A.3: Shares of real prices and on-sale prices by length of price spell



Shares of observations that are to-the-cent prices and on-sale by spell-length for the retailer data (corresponding to column (3) in Table 1.1). A product-week-store observation is allocated to spell length T if the price on that week is identical across T consecutive weeks. The light yellow line is the CDF of observations in the data. The dark circles are the share of observations, conditional on spell length, that are “real” (i.e., prices ending to-the-cent). The lighter triangles are the share of observations, conditional on spell length, that are on-sale (i.e. net price is less than the regular price).

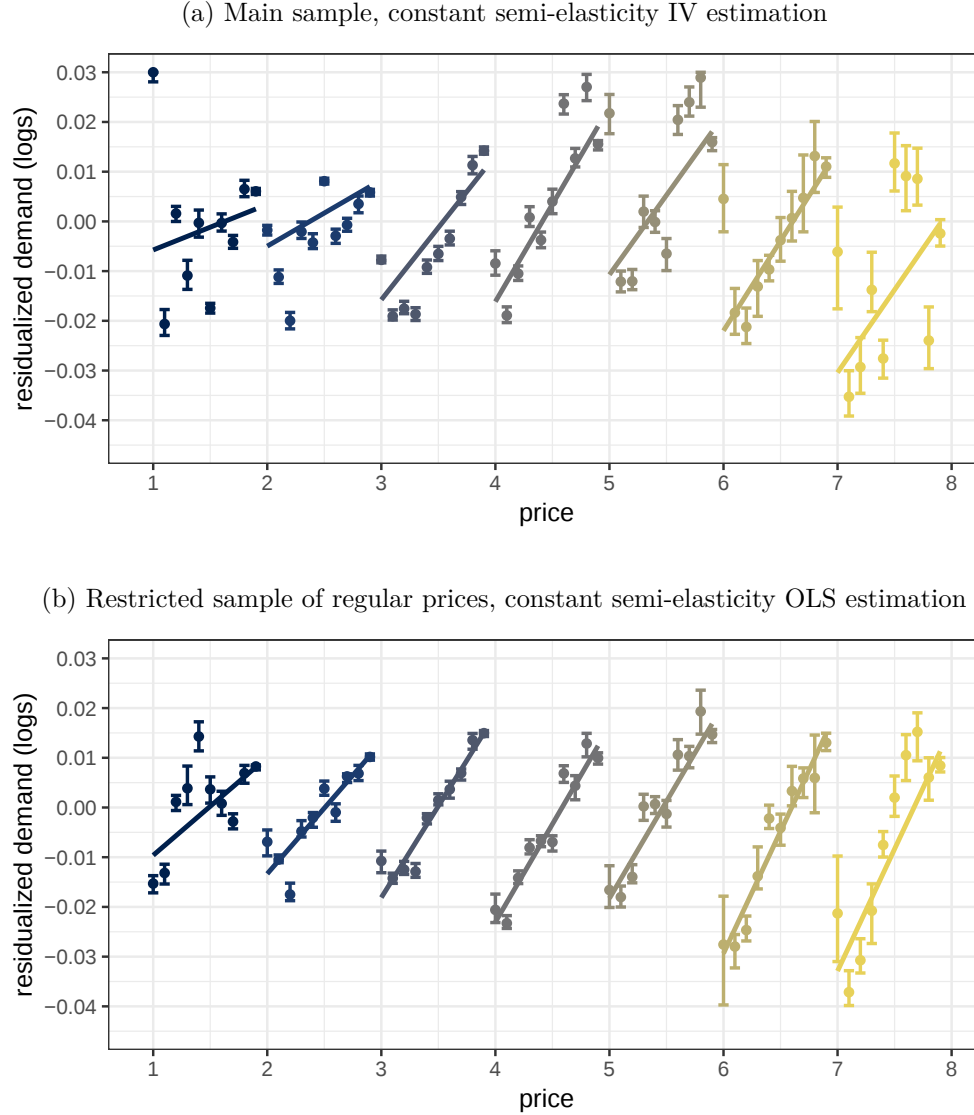
Figure A.4: Retailer Robustness: Residualized demand of 12 products and 70 retailers using Nielsen data



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

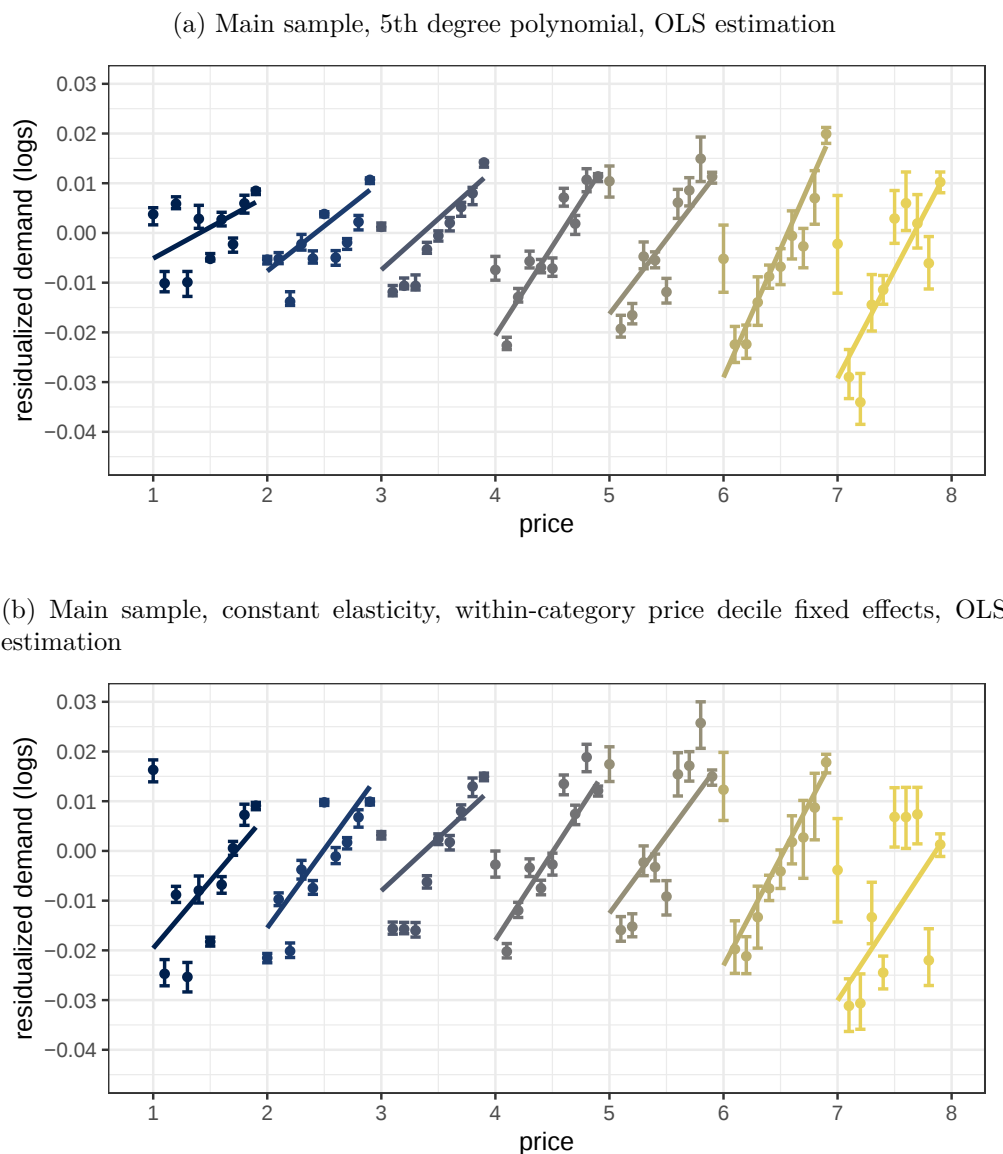
The top panel is using data on all observations of at least 2-week spells (Column (2) in Table A.1). The bottom panel presents the results from a restricted sample of at least 6-week spells (column (3) of Table A.1). Data is truncated for prices between \$1 and \$9.99 for clarity. For further details see Section 1.4 for the empirical specification and Appendix A.4 for data description.

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



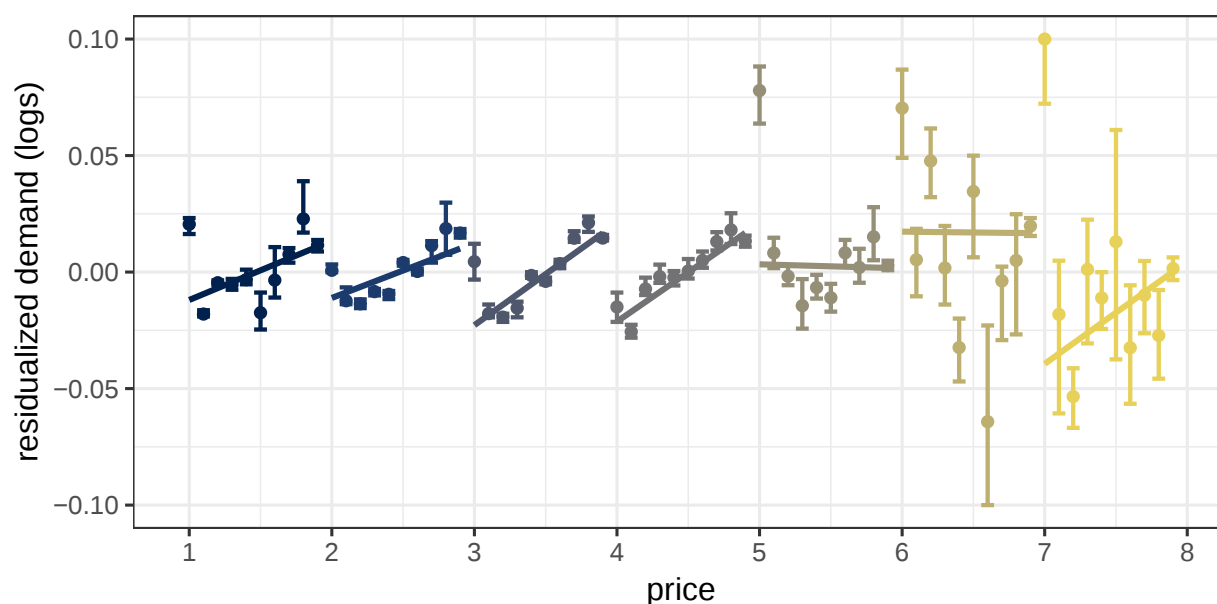
Robustness: Alternative demand structure. Net log-demand by price, residualizing product-level price *semi*-elasticity, cross-*semi*-elasticity, seasonality, store fixed effects, and promotion effects. Regressions are conducted at the product level (1710 products, 248 stores, 177 weeks). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Fixed effects are estimated at 10-cent bins, and 95% confidence intervals are result of 200 cluster-bootstraps by store. The top panel is using data on all to-the-cent prices from a national retailer data (Column (3) in Table 1.1), and the regression is run with instrumental variables, as in Specification 1.6 but on prices rather than log-prices. 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 panel presents the results of an OLS constructed on regular prices only (column (4) of Table 1.1). Data is truncated for prices between \$1 and \$7.99 for clarity. For further details see Section 1.4.

Figure A.6: Demand structure robustness: Residualized polynomial demand and rank fixed-effects



Robustness: Alternative demand structure continued. Net log-demand by price. Top panel: residualizing product-level 5th degree polynomial in prices, cross-elasticity, seasonality, store fixed effects, and promotion effects. Bottom panel: residualizing product-level log-price, the decile of product price within the concurrent price distribution of items from the same product category, cross-elasticity, seasonality, store fixed effects, and promotion effects. Regressions are conducted at the product level (1710 products, 248 stores, 177 weeks). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Fixed effects are estimated at 10-cent bins, and 95% confidence intervals are result of 200 cluster-bootstraps by store.

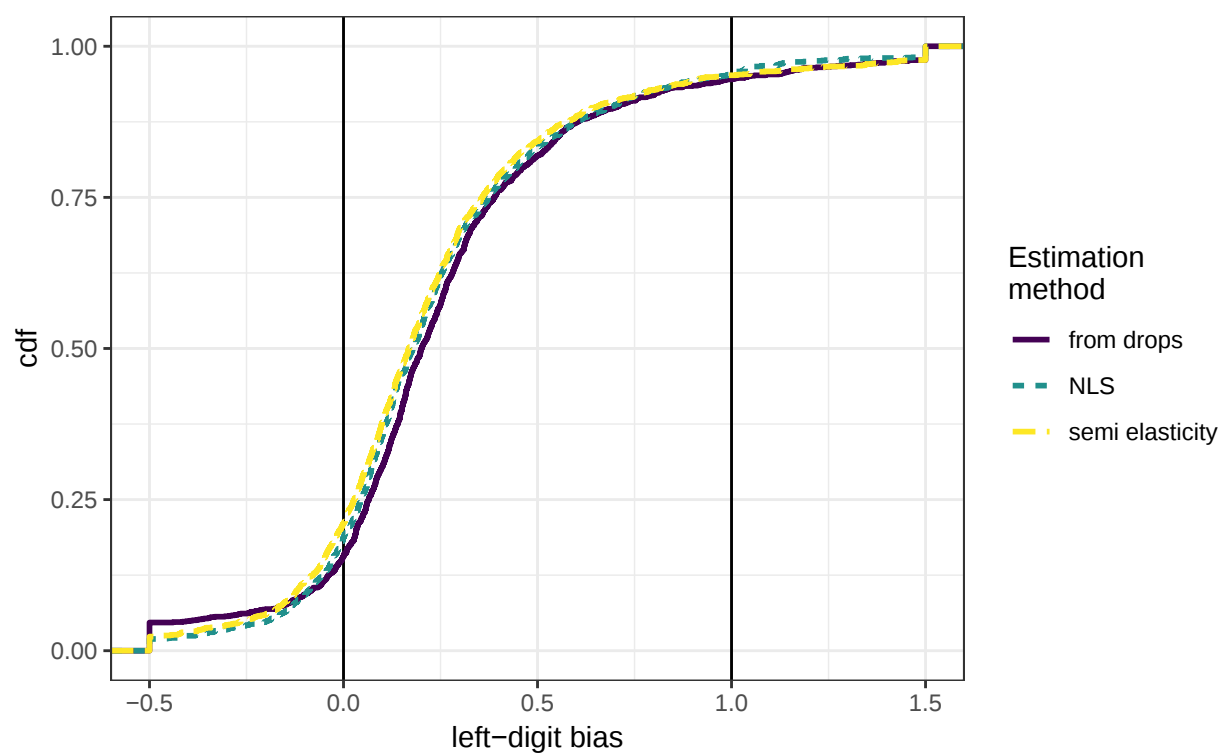
Figure A.7: Sample robustness: Residualized demand for an alternative sample of products



Robustness: alternative sample. Net log-demand by price, residualizing product-level price elasticity, seasonality, store fixed effects, and promotion effects. Regressions are conducted at the product level (2149 products). Different dollar digits are designated by color. Within-dollar linear fits, weighted by number of observations, are added in solid lines. Fixed effects are estimated at 10-cent bins, and 95% confidence intervals are result of 100 cluster-bootstraps by store. Top panel shows the entire sample of to-the-cent prices, while the bottom restricts attention to regular prices only.

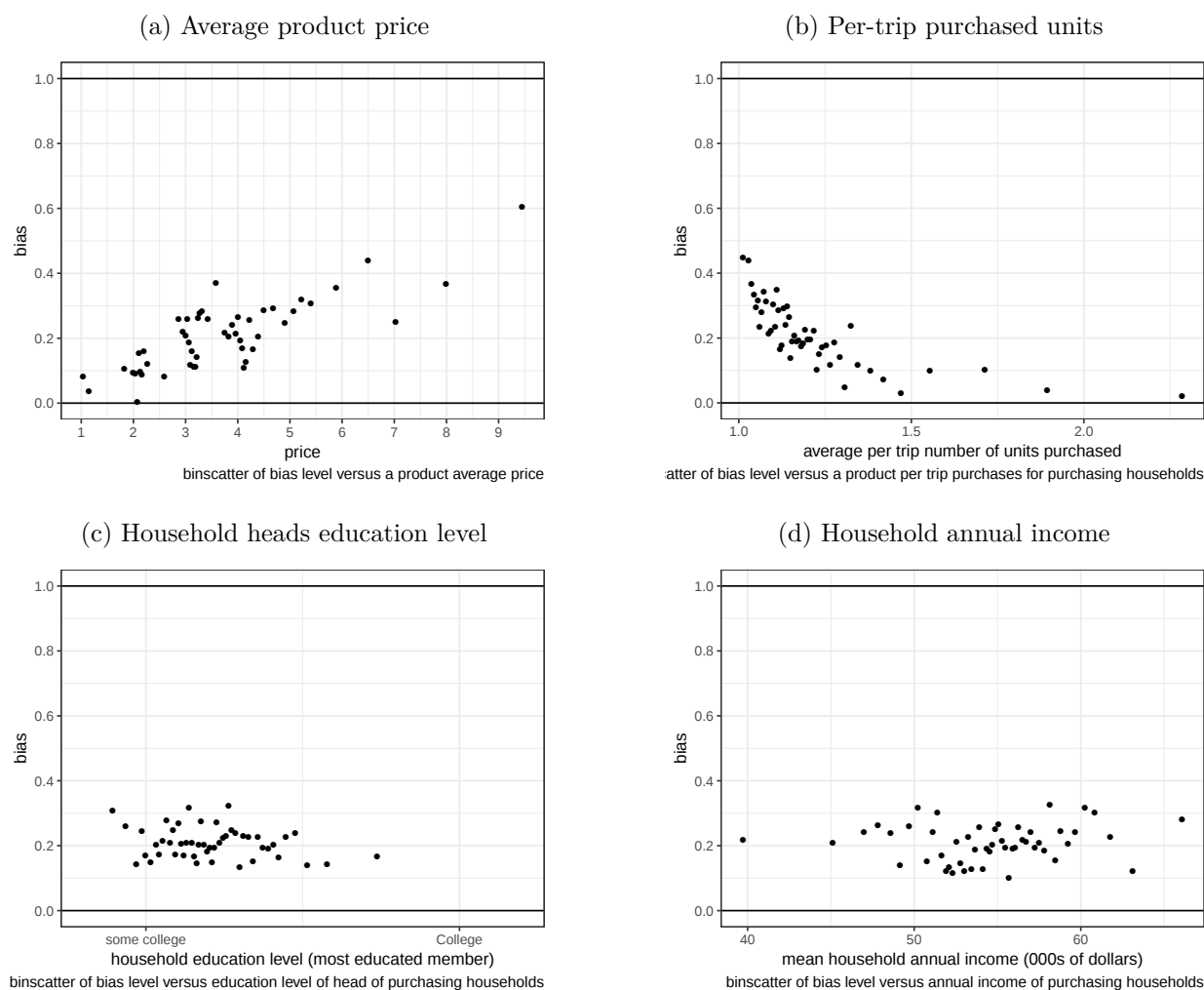
The sample is of 2149 products of high availability but also highly concentrated prices (125 with more than 75% of observations at 99-ending prices, the rest with more than 75% of observations sharing common dollar digit). Sample consisted of 55.5 Million observations spanning the same stores and period of time as for the main sample. Data cover \$292 Million in annual revenues via sale of 170 Million units per year. Mean price of these products is \$2.20 (1.27), with 21.7% of observations representing on-sale prices.

Figure A.8: CDF of left-digit bias estimates on Retailer data, three estimation methods.



The figure shows the CDF of product-level left-digit bias estimates, corresponding to panel B of Table 1.2. The dark solid line, “From drops”, is the CDF of left-digit bias estimates at the product level using the drops in demand at dollar crossings coupled with elasticity of demand and the dollar digit. Dashed green, “NLS”, are estimates from non-linear least squares. Light long dashed, “semi elasticity”, are the estimates from a log-linear regression with constant semi-elasticity of demand. For details see Section 1.4.

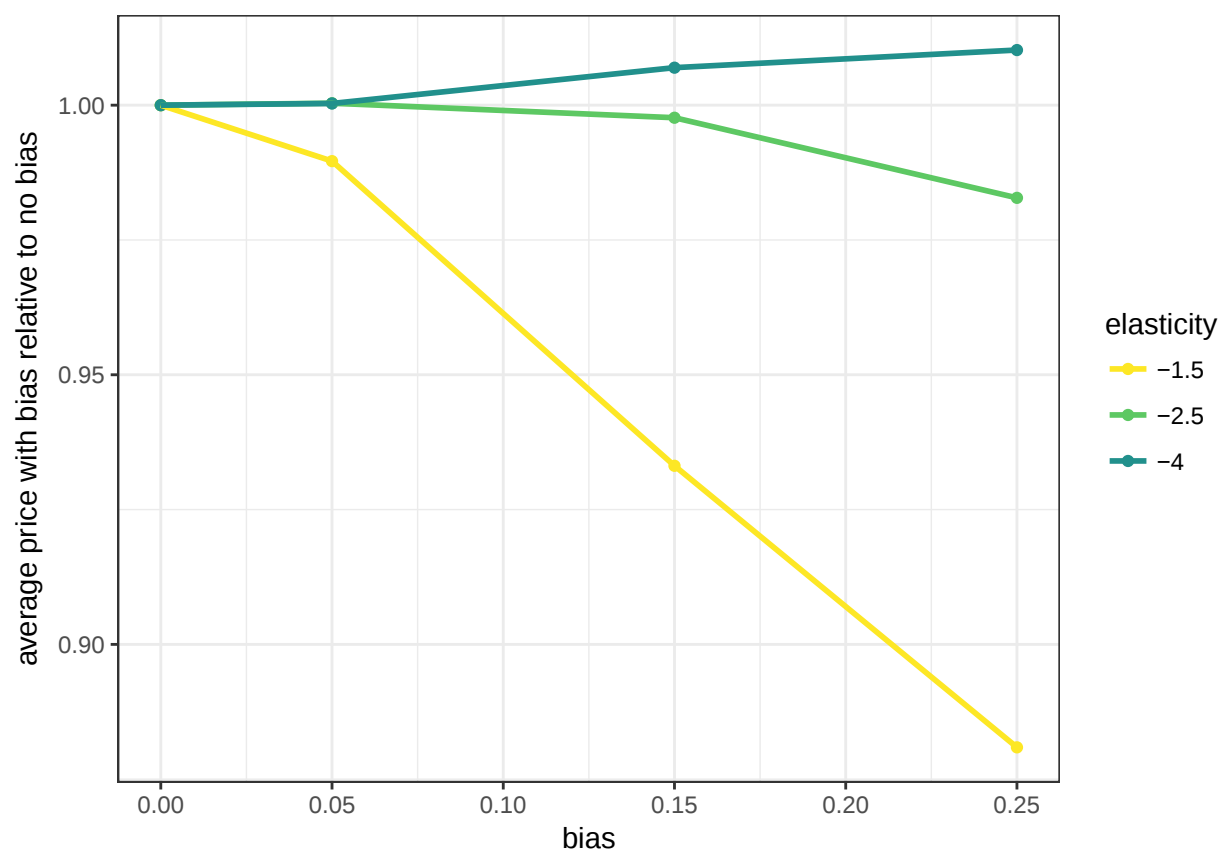
Figure A.9: Bias Heterogeneity: Correlates of product and clientele characteristics with left-digit bias



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

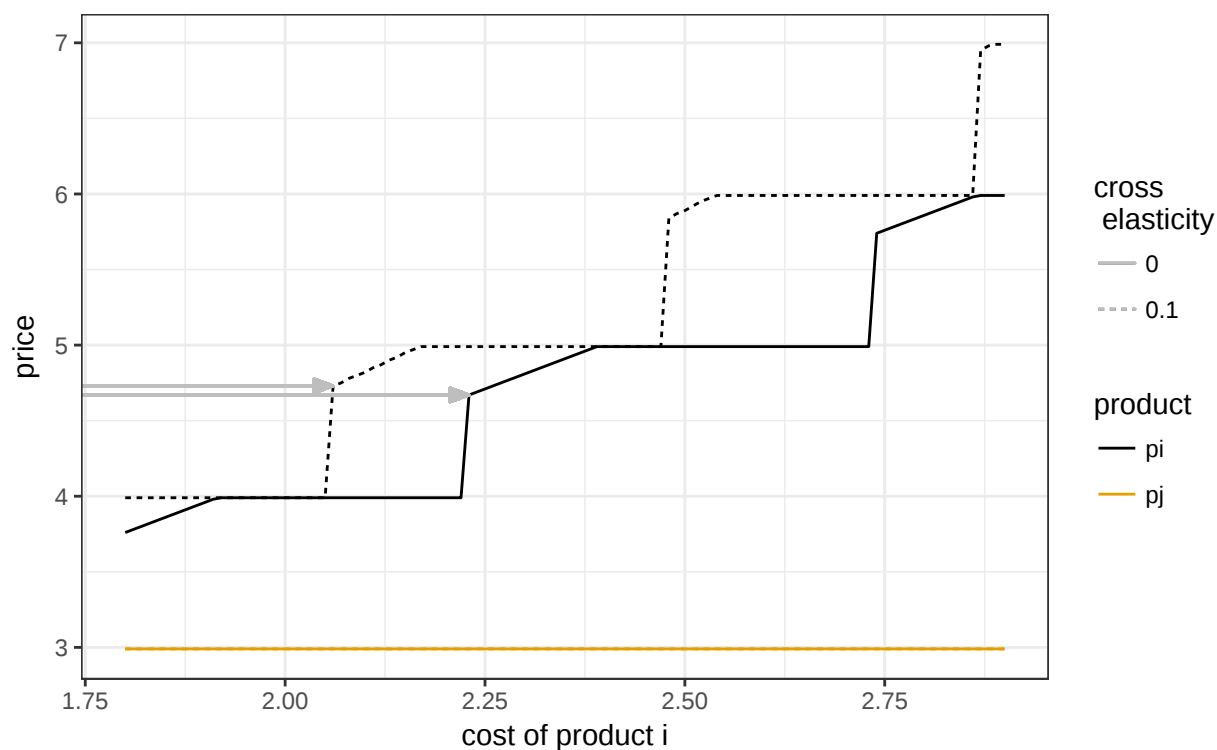
Panels (a)-(d) show the association with the product overall average price, the average number of units a purchasing household purchase of that item in a single trip, the education level of purchasing household heads (maximum level if two heads of household), and the annual income of purchasing households. For further details see Appendix Section A.5.

Figure A.10: Simulation: Relative price level by bias



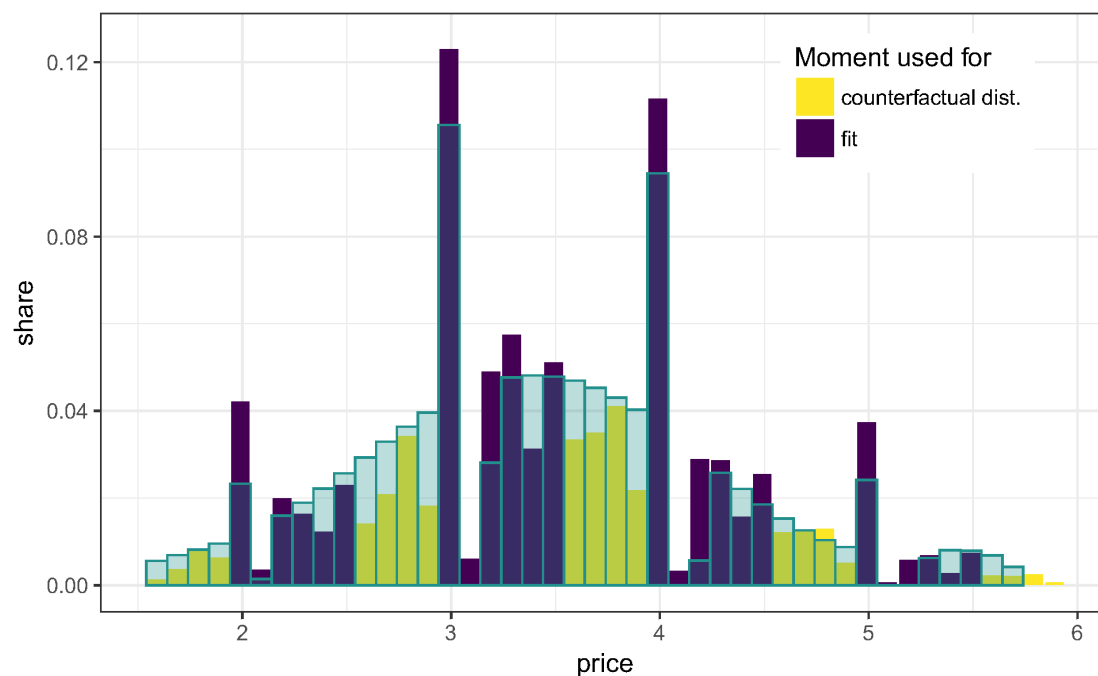
The figure shows the the average price of products set by an optimizing firm by the level of left-digit bias, keeping everything else fixed. The horizontal axis is the degree of left-digit bias, the vertical axis is the relative price to a no-bias case, normalized to 1. Each curve represents a different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

Figure A.11: Illustration: pricing schedule of a 2-product monopolist



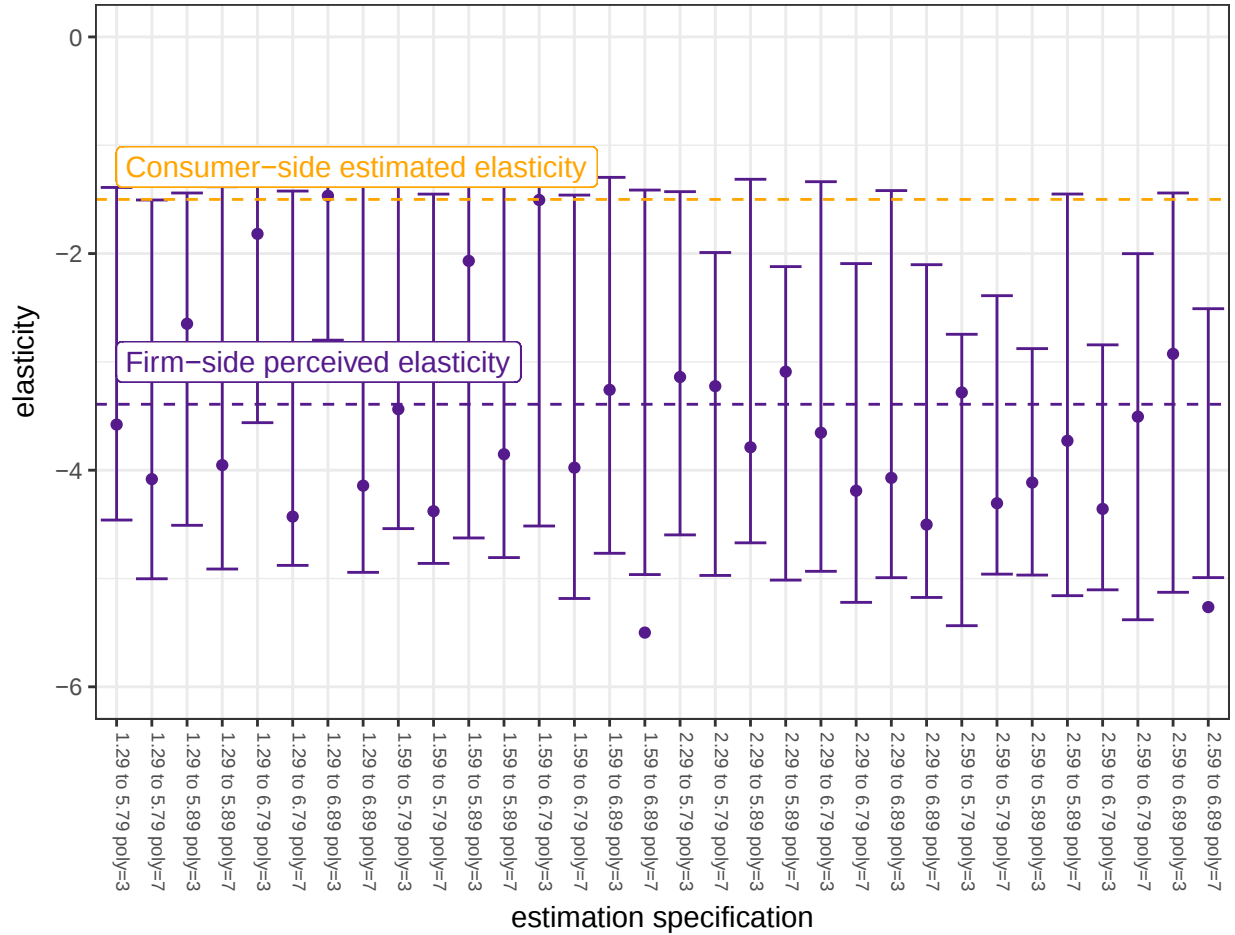
The figure gives an example of a pricing schedule of two products with and without cross elasticities of substitution between the items. The same-price elasticity of each product is -2. Black curves are the prices of product i and orange curves, which lie on top of another, are the prices of product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

Figure A.12: Minimum distance estimation: empirical and predicted moments



The figure shows the empirical moments and predicted moments from a minimum distance estimation. Each moment is share of prices in a 10-cent bin. The empirical moments are the shares of regular prices for the 1710 products, shown in solid bars. Dark bars are the moments used for fitting, while the light bars are excluded from the fit and used to create the counterfactual price distribution. The predicted moments are shown in semi-transparent light green. For further details see Section 1.5.

Figure A.13: Estimated firm-perceived elasticity



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

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

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

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

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

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

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

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

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

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

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

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

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

Table A.4: Direction of welfare effects under various assumptions

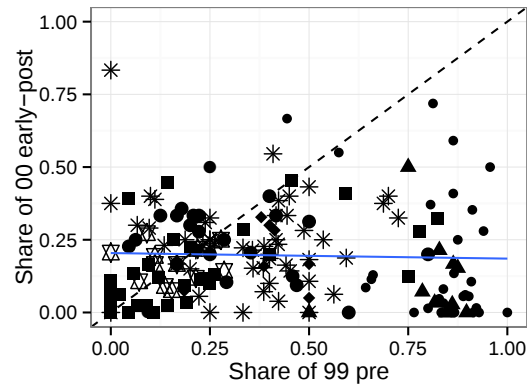
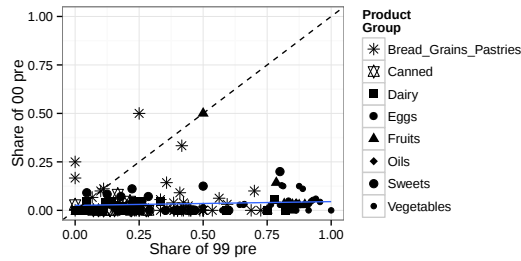
Value dist.	effect of...	$\Delta = 0$				$\Delta = 0.99$			
		CS	DWL	P	Π	CS	DWL	P	Π
Beta	$\theta > 0 \leftrightarrow \theta = 0$	-	-	+	+	-	+	+	-
$\sim 8 * \beta(1, 3)$	$\hat{\theta} < \theta \leftrightarrow \hat{\theta} = \theta$	+	-	-	-	+	-	-	-
Exponential	$\theta > 0 \leftrightarrow \theta = 0$	?	-	?	+	?	+	?	-
$\sim 1 + 3 * \exp(1)$	$\hat{\theta} < \theta \leftrightarrow \hat{\theta} = \theta$	?	+	?	-	?	+	?	-
Gamma	$\theta > 0 \leftrightarrow \theta = 0$	-	-	+	+	+	-	-	-
$\sim \Gamma(8, 0.5)$	$\hat{\theta} < \theta \leftrightarrow \hat{\theta} = \theta$	+	+	-	-	-	+	+	-
Normal	$\theta > 0 \leftrightarrow \theta = 0$	-	+	+	+	?	+	+	-
$\sim N(4, 3)$	$\hat{\theta} < \theta \leftrightarrow \hat{\theta} = \theta$	+	-	-	-	?	?	?	-
Pareto	$\theta > 0 \leftrightarrow \theta = 0$	?	-	?	+	-	+	+	-
$\sim 5 * \text{Par}(3) - 5$	$\hat{\theta} < \theta \leftrightarrow \hat{\theta} = \theta$	?	+	?	-	+	?	-	-
Uniform	$\theta > 0 \leftrightarrow \theta = 0$	-	+	+	+	-	+	+	-
$\sim U[0, 6]$	$\hat{\theta} < \theta \leftrightarrow \hat{\theta} = \theta$	+	-	-	-	+	-	-	-

Each cell in the table describes the direction of an effect (as described on the second column) on one of the measures (column title - Consumer Surplus, Dead Weight Loss, Price, Profits), given Δ (top row) and value distribution (left-most column), averaged across a cost distribution as in Section 1.6. “+” symbolizes higher, “-” lower, and “?” means that the effect changes signs for different values of θ . For example, the “-” in the bottom right cell means that when firms underestimate the bias ($\hat{\theta} < \theta$), consumers valuation of the item is distributed uniformly between 0 and 6, and consumers’ focal price ending is 99, then DWL is lower than if firms were to perceive the bias correctly ($\hat{\theta} = \theta$). Across rows, and regardless of Δ , lost profits from underestimation are in the range of 1%-2%.

Appendix B

Appendix to Chapter 2

(a) Correlation between share of 99 pre and 00 pre (3-10/2013) (b) Correlation between share of 99 pre and 00 immediately after policy change (1-2/2014)



(c) Correlation between share of 99 pre and 00 a year after policy change (1-2/2015)

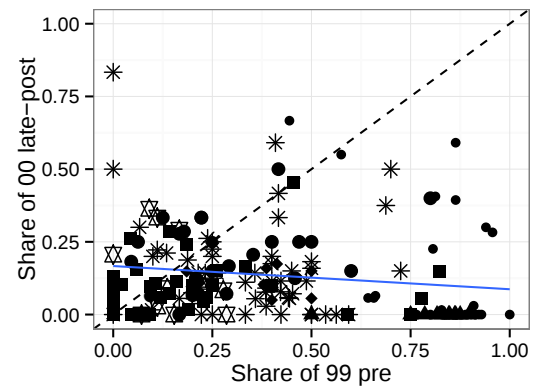


Figure B.1: Correlation between pre-shares of 99 and post-shares of 90 and 0. Each dot is a product in the CBS data.

The figures show that shares of 00-ending prices at different periods against 99-ending prices at the pre-period, by product-type in the CBS data. The shares of 00-ending prices per product are calculated before the reform announcement, at the two months following the reform, and at early 2015 a year after the reform.