

Learning and Limitations of Heuristic Pricing: Evidence from a Reform

Avner Strulov-Shlain*

October 25, 2026

Abstract

Pricing is a core firm activity, yet it is unclear how firms set prices. Do they maximize a demand model, or satisfice and follow heuristics held to a good-enough standard? The two are hard to separate, because a rule of thumb can often be relabeled as optimization net of the cost of doing better. An Israeli reform separates them: it banned prices ending in non-existent denominations (like .99), forcing supermarkets to re-price. Within a class of left-digit-bias demand models—downward sloping, with discontinuities at round numbers—a firm optimizing even under parameter uncertainty would move to the next-best ending, .90, and avoid the dominated .00. Instead, about 20% of prices moved to .00. Which prices moved there was idiosyncratic, and the move reversed to .90 only over the following year, paced by the shortfall each firm could detect against its own past sales. Demand itself was economically stable across the reform on several tests. The evidence favors firms that knew .99 worked without having a model that ranked unfamiliar endings. Forced exploration eventually brought profits closer to optimum than they were pre-reform. Thus, heuristics can bring firms close to the optimum, but they require learning what is possible and can be fragile when conditions change.

JEL Codes: D22, D83, L11, L81, M31

Keywords: pricing, heuristics, learning, left-digit bias, firm behavior, behavioral firms, satisficing

*Strulov-Shlain: University of Chicago Booth School of Business, avner.strulov-shlain@chicagobooth.edu.

I am thankful to Stefano DellaVigna for endless encouragement and support, and to Jean-Pierre Dube for honest discussions and detailed feedback. Ned Augenblick, Zarek Brot-Goldberg, Ben Handel, and Dmitry Taubinsky provided many thoughtful comments. Christy Kang provided excellent research assistance. This research greatly improved thanks to discussions with Nick Barberis, Modibo Camara, Ben Enke, Alex Frankel, Wes Hartmann, Johannes Hermle, Ella Honka, Ali Hortacsu, Alex Imas, Avi Goldfarb, Tesary Lin, Sanjog Misra, Sarah Moshary, Muriel Niederle, Brad Shapiro, and Richard Thaler. All errors are my own. I am indebted to Itai Ater, Roni Rojkin (ICC), Saul Lach, and Merav Yiftach (CBS) for greatly assisting me with data collection and thus making this project viable.

“There was a 2%–3% decrease in revenues in Jan–Feb . . . No one can explain why that happens . . . I’m in business for 25 years, and had never seen such months.” (CEO of Mega supermarket chain, March 2014)

1 Introduction

Learning to set prices is important but complicated, as the opening quote illustrates. The standard assumption in economics is that firms optimize a known objective function: they understand demand and choose prices accordingly. A long tradition doubts this, arguing that firms satisfice—following rules of thumb until performance is good enough (Simon, 1955; Romer, 2006). The two views are hard to tell apart, because any satisficing rule can be redescribed as optimization net of the cost of doing better, whether that cost is computational, informational, or attentional. They are also hard to tell apart because in a stable environment they largely agree: a rule refined by experience and a solved demand model can prescribe nearly the same price. They come apart when conditions change, since a firm that never knew demand cannot simply re-derive it when its rule stops working, and predictions built on the standard assumption become unreliable there. Separating these explanations is thus important both for the decision-making firm and for researchers studying what will happen when conditions change. It requires a specific shock that this paper uses.

The separating setting is a reform that restricted available price endings. In 2013, Israel banned prices ending in denominations of coins that no longer exist, which removed the .99 ending¹ that supermarkets had relied on for decades. The reform tells the two views apart because they predict different behavior once .99 is gone. Within the demand-model class studied here, a firm that optimizes even under parameter uncertainty would move to the next-best ending, .90, and avoid the dominated .00; a firm that had learned .99 works without learning why may have no way to rank the untried endings until it tries them.² The empirical evidence favors the second description: about 20% of prices moved to the dominated .00s, and adjustment away from .00 took almost a year.

Under the standard view, *model-based optimization*, firms maintain a model of demand and choose prices to maximize profits given that model. I use the term for a procedure that maximizes expected profits using a structured demand representation that links candidate prices. What defines the view is the *max* operator: whatever the firm believes, it prices where expected profits are highest, and updating one belief re-ranks all candidate prices at once. The model may be held with uncertainty, and acting on it may be costly, so behavior can be optimal up to information and

¹For brevity I will write .XX, such as .99, to designate .XX ending prices. e.g., .99 means prices like 2.99 or 5.99 and .00 are prices like 2.00 or 7.00.

²The relevant version of the model-based view restricts demand to be downward sloping, with potential discontinuities at round numbers. An unrestricted model—including one with an independent demand intercept for every ending—can rationalize any behavior, including a move to .00, but then provides no separating prediction; Section 4.2 treats this surviving alternative explicitly.

attention frictions (Hortaçsu et al., 2024; Sims, 2003; Gabaix, 2014). Importantly, a useful model creates a structure that links candidate prices to one another.

Under the alternative, which I call *heuristic pricing*, firms follow a rule (e.g., “price just below the round number”) and check whether it meets a standard of good enough.³ Because the optimum is not observable, the standard is built from benchmarks the firm can actually see such as how the product used to sell, or how it sells at the firm’s other stores. Throughout, I use *benchmark* for an observable comparison of this kind and *aspiration* for the performance threshold applied to it; the data measure the precision of the benchmark, not the aspiration cutoff itself. This is also how firms are observed to form aspirations—they track their own past performance and turn to comparisons with rivals mainly when performance falls short (Cyert and March, 1963; Lant, 1992; Blettner et al., 2015; Börgers and Sarin, 2000).

Heuristic pricing, as used here, updates action rules from local feedback without such cross-action generalization. Because the comparison evaluates only the action taken, learning is local: it can tell the firm whether its current rule clears the benchmark, but says nothing about actions never tried.⁴ This is what makes exploration necessary: what counts as “good enough” is bounded by what the firm has already tried, so the standard can improve only when the firm tries or observes something better. Firms therefore choose actions and improve until performance clears an accessible benchmark, then stop. This is satisficing in Simon’s original sense.

To illustrate the core identification idea of heuristics versus model-based optimization, consider a thought experiment. Imagine three options with an objective ranking of $a \succ b \succ c$. In steady state, the decision maker chooses a , consistent with the ranking. Now suppose option a is removed from the choice set. If the decision maker had a model informing the full ranking, they would immediately switch to b . But if they had arrived at a adaptively—learning that a works well without understanding *why* or how well alternatives work—they may not know that $b \succ c$. In that case, they might initially try c , and only through experimentation learn to switch to b . Trying c after seemingly knowing it is the worse option reveals the lack of a model.⁵

The reform studied in this paper mirrors this thought experiment. In 2013, the Israeli government banned pricing at the level of coins that no longer existed (the equivalents of the U.S. penny and nickel), forcing companies to use prices ending in .X0 (Ater and Gerlitz, 2017). For example, 2.96 was banned, while 2.90 or 3.10 were allowed. Before the reform, approximately 45% of supermarket

³Heuristics and satisficing are often treated as rivals, but they answer different questions (Simon, 1955; Artinger et al., 2022): a heuristic is a statement about *actions* (“price just below the round number”), satisficing a statement about *standards* or aspirations of what counts as good enough (“we should make this much in sales”). In Simon’s original formulation both elements were present—an aspiration level and search among alternatives—and later literature split them apart and studied each separately against optimization (Artinger et al., 2022).

⁴Cost-plus pricing, for example, belongs to the heuristic view if the markup is adjusted through trial and error against realized performance, and to the model-based view if the markup is derived from beliefs about demand. The line between the views is not the firm’s optimization budget but whether a demand model ranks the alternatives.

⁵This is akin to Caplin and Dean (2015) NIAS necessary condition for “any arbitrary finite dataset to be consistent with a model of costly information acquisition”: a rationally inattentive decision maker cannot choose an option it knows to be worse than another. Failing this condition implies the data are inconsistent with models of costly information acquisition (also known as Rational Inattention).

prices ended in .99 (e.g., 2.99), while .00 prices were rare. This distribution is consistent with pricing that accounts for left-digit biased demand⁶—a demand structure driven by consumers’ distorted price perception, which leads to discontinuous drops in demand at round prices. Under such a demand structure, prices that would have ended with low digits in the absence of bias (e.g., 5.00 or 5.29) are better set with a .99 (e.g., 4.99). In the notation of the thought experiment, .99 prices represent option a , .90 prices option b , and .00 prices option c , with $a \succ b \succ c$.

When .99s were banned, model-based pricing predictions still implied that .00 prices should be avoided. The magnitude of left-digit bias that made .99s highly attractive also makes .90s dominate .00s, even though the trade-off between .90 and .00 is stronger.⁷ If firms understood the demand structure, they would have immediately shifted to .90s. Instead, at the time of the reform, 20% of prices ended in .00. In the years that followed, the share of .00s declined as firms gained post-reform experience. Products whose prices shifted from .99 to .90 tended to remain at .90, while those that shifted to .00 typically reverted to .90 over time. About a year after the reform, firms had largely converged to .90.

The initial transition to .00 has little structure in observed covariates or in additive chain and store propensities, but strong product-specific structure. Pre-reform prices, the reliance on .99s, chain size, local competition, and chain and store identities all explain little of which product-stores first moved to .00 rather than to .90. The regularity is organized *between products* and is mostly unobserved: a variance decomposition attributes the bulk of the explained variation to product identity, robustly so out-of-sample, with store and chain identity explaining little.⁸ Products with higher pre-reform shares of .99s initially rounded less to .00s, but this too is a between-product regularity: within a product, stores that leaned more on .99 were no less likely to round to .00. These idiosyncrasies are telling about a lack of knowledge on how to react to the reform.

In contrast, systematic patterns appear in the correction away from .00 rather than in the initial move and are informative about the learning process. Product-stores that changed originally to .90 tend to stay at .90, while those that changed to .00 switch to .90, a pattern consistent with feedback-driven correction. Accordingly, a balanced back-and-forth between the price endings is rare (what a pre-designed A/B test would look like). Some chains move out of .00 in stepwise fashion rather than smoothly, consistent with chain-level policy changes rather than gradual store-by-store updating. In addition, product-stores experience faster reversion when the demand they realize at .00 falls further below that of same-chain peers still priced at .90. This reversion most clearly happens where the firm’s own pricing history provides a precise benchmark at .99 against

⁶That supermarkets recognize they face left-digit biased demand, to some degree, was also documented in U.S. supermarkets (Strulov-Shlain, 2023).

⁷There is indeed a trade-off introduced by the reform: pricing below the round price threshold (e.g., 4.90 vs. 4.99) incurs a larger revenue loss, and the discontinuity becomes “smoother.” Yet, left-digit bias that justified many .99s also supports the use of .90s, but not .00s.

⁸A Shapley-Owen decomposition attributes 55% of the explained variance in the first-transition outcome to product identity, 28% to store identity, and 11% to chain identity; in a cross-validated version the store share falls to 4% while the product share rises to 75%.

which to detect the shortfall. These patterns are consistent with feedback-driven adaptive learning, coordinated at the chain level and paced by the strength of each product’s prior just-below heuristic and subsequent demand changes.

One alternative explanation is that demand changed due to media and consumer attention, and left-digit bias became temporarily less important, thus justifying the move to .00. Three complementary tests do not support it (although they cannot disprove every such change in demand): the elasticity is economically stable, the .90-versus-.00 gap is similar across the two halves of 2014, and a pre-reform demand model projects forward with only modest drift.⁹ Across instrumental-variable strategies, the preferred specification estimates demand at .00 to be at least 6.5% lower than at .90, on 17 scanner products and in the post-reform periods where .00 is common enough to be informative. I therefore regard a reform-induced reversal in the demand ranking as possible but less likely than stable left-digit biased demand. Section 4 works through other alternatives one by one, and Table A4 summarizes them. None of the alternatives is logically impossible, but each fits the transition less well or requires more auxiliary assumptions than satisficing. For example, full knowledge within the specified demand class does not explain why firms used .00, and complete ignorance does not explain why they used .99 before the reform and .90 after.

A simple calibration quantifies possible profit consequences. Using estimated demand parameters and observed price distributions, I compare actual pricing to a counterfactual of full optimization. Before the reform, estimated forgone gross profits were about 0.6%. At the reform’s onset they rose to about 0.8%, then declined to about 0.5% by the end of 2014, below the pre-reform estimate. The reform-induced transition therefore appears to have improved pricing relative to the pre-reform benchmark, while leaving firms short of the frictionless optimum. The levels (0.6% is on the conservative end) depend on demand and cost assumptions; the spike and subsequent decline below baseline hold across the demand specifications and parameter grid examined.

This paper does *not* claim that heuristic pricing is necessarily suboptimal. Estimating demand is costly, as is figuring out the optimal action given demand, or even what an optimal profit might be, so rules of thumb held to a good-enough standard may be a reasonable response. The point is descriptive: firms *can* find not-terrible prices without having to know the objective function, and at the same time that can leave them short of the frictionless optimum. This has implications for both researchers and practitioners. For researchers, positive economics may describe steady-state behavior reasonably well even if not perfectly, but predictions about responses to shocks require caution. For practitioners, it highlights the value of exploration and the risks of relying on heuristics when conditions change or when the idea of what is possible (and hence the stopping rule) is vague.

The main contribution is to show that near-optimal steady-state pricing need not rest on knowledge of the optimum. Firms that had priced well for decades could not rank the endings left to them when .99 was banned, and found the better one only by trying the worse one. The partial

⁹The average price elasticity is stable over time, at about -2 on average.

correction took about a year and, within the model, left forgone profits below their pre-reform level. This paper builds on Strulov-Shlain (2023), which estimated significant left-digit bias and persistent misoptimization for US supermarkets, in order to learn about the nature of misoptimization and learning.

The paper advances the satisficing literature. Within it, aspiration-based rules have mostly been argued for either by procedural evidence or based on model fit (Artinger et al., 2022). For example, in pricing, used-car dealers set an aspired price and mark it down at fixed intervals (Artinger and Gigerenzer, 2016). A recent paper (Salant and Spenkuch, 2026) provides the cleanest empirical test for satisficing over maximization from comparative statics in how noisily chess players evaluate moves of differing complexity, holding choice sets fixed. The reform provides a complementary identification: what an aspiration-based pricing rule does when its action is outlawed (Section 4.1 maps the reform onto their framework). That literature leaves open how aspiration levels are formed and adjusted. The current paper identifies a measurable benchmark as the product’s own sales history, and its precision further governs the speed of correction.

Related work on firm learning has studied firms confronting a changed environment: learning about competitors’ play (Doraszelski et al., 2018), about a new demand structure (Huang et al., 2022), or about macroeconomic and regulatory shifts. The reform differs from settings in which firms must learn a new demand structure. It reduced the feasible set while leaving the estimated demand relationship economically stable. A firm with a structured demand model would therefore face a relatively direct re-ranking problem, yet aggregate adjustment took about a year.

The dynamic itself is not specific to price endings. A shock reveals that a firm does not know its own demand, and the firm turns to experimentation: ZipRecruiter ran a first demand experiment to learn it was leaving dozens of percent of profit on the table (Dubé and Misra, 2023); a Mexican payments firm randomized merchant fees after a competitor’s undercut revealed it did not know its own fee elasticity (Gertler et al., 2025); a Venezuelan lender experimented with credit limits once its heuristics no longer fit market conditions; and Philippine banks did the same under a regulator-mandated uniform digital-payment price.¹⁰ In each case above a firm initiated an experiment. The reform instead forced exploration on firms that did not recognize the problem, which allows us to learn about both their prior knowledge and learning process.

This paper contributes to the literature on firm misoptimization by focusing on the source of it. A growing body of work documents persistent deviations from profit maximization (Bloom and Van Reenen, 2007; Cho and Rust, 2010; DellaVigna and Gentzkow, 2019; Goldfarb and Xiao, 2011; Gertler et al., 2025; Hortaçsu et al., 2024; Huang, 2021; Salant and Spenkuch, 2026; Shapiro et al., 2021; Strulov-Shlain, 2023). Further, whether firms maximize at all is a long-standing question (Romer, 2006). One relevant pattern is that firms set prices and wages in coarse units rather than

¹⁰The Venezuelan and Philippine cases are documented in work in progress by Higgins and Kluender (credit limits at the buy-now-pay-later provider Cashea) and by Gonzales, Higgins, Tóth, Tsai, and Wang (bank pricing of digital payments in the Philippines); drafts had not circulated at the time of writing.

fine ones (Aparicio and Rigobon, 2023; Dube et al., 2017; Reyes, 2024)—the phenomenon this paper studies at the level of price endings where it is justified. These findings raise a puzzle: how can firms persistently fail to optimize when the stakes are high? One answer is that firms face constraints that prevent optimization. Another, suggested here, is that firms may not actually know the objective function and have to rely on a benchmark. They arrive at reasonable strategies through trial and error, but may lack the knowledge to fully solve more complicated problems, to know what the feasible optimum is, and to immediately adapt when conditions change.¹¹ This interpretation provides a parsimonious explanation for both the steady-state behavior (potentially close to but not at optimal) and the transition dynamics (slow adjustment with initial mistakes).

The paper proceeds as follows. Section 2 describes the setting, data, and left-digit biased demand. Section 3 presents the results on pricing patterns and demand response. Section 4 interprets the findings. Section 5 discusses implications.

2 Setting and Data

2.1 Reform details

On October 17, 2013, the Israeli Ministry of Economics announced that, starting January 1, 2014, it would begin enforcing an existing law banning consumer pricing in denominations of coins that no longer exist.¹² Israel had abolished the 1-Agora and 5-Agorot coins (the Shekel consists of 100 Agorot) in 1991 and 2008, respectively.¹³ Yet prior to 2014, the majority of prices ended in denominations different from 10-Agorot multiples—most commonly with 9- and 5-endings. Therefore, most prices had to change due to the reform.

This policy was motivated by consumer concerns that they were being misled by non-existent prices (e.g., a posted price of 4.99 that is actually rounded to 5.00 at checkout; prices in Israel are VAT-inclusive). Consumers either paid the total basket price electronically or the rounded price (to the nearest 10 Agorot) when paying in cash.

The reform was covered in the national press, at the announcement in October 2013 and again around enactment in January 2014, framed as a consumer-protection measure against misleading price endings¹⁴; the opening quote, from a March 2014 interview with the CEO of the second-largest

¹¹This is corroborated anecdotally by the reasons firms wished to collaborate with academics in the cases of Gertler et al. (2025) and Dubé and Misra (2023).

¹²With the exception of “continuous” products—gas, water, and electricity. The rules also do not apply to business-to-business pricing.

¹³<https://web.archive.org/web/20201127080757/https://www.boi.org.il/en/Currency/CurrentCurrencySeries/Pages/Default.aspx#Top>

¹⁴e.g., see some examples of price changes in this article from December 2013 (in Hebrew) <https://web.archive.org/web/20131220045135/ynet.co.il/articles/0,7340,L-4466463,00.html>, which works through examples of the price changes. The examples it presents as objectionable are round-ups: a 5.99 price becoming 6.00 is described as firms using the reform to raise prices.

chain, is one example of the coverage that followed. Whether this publicity temporarily changed how consumers respond to price endings is treated as an empirical question rather than an assumption: Section 3 tests demand stability around the reform directly.

2.2 Data

I use supermarket pricing data to learn about the pricing dynamics, and scanner data to learn about demand. Data are drawn from three sources: two are shelf price datasets from before and after the policy change, and the third is scanner data containing revenues and quantities sold. In all datasets, the unit of observation is a product in a store in a given period, though the sampling period differs by source—every 2–4 weeks in the ICC data, monthly in the CBS data, and weekly in the scanner data.

Pricing data. The main pricing dataset was collected by the Israel Consumer Council (ICC), a consumer advocacy group. ICC employees visited supermarket stores every 2–4 weeks between early 2013 and early 2015 and recorded the prices of a predetermined list of products, meant to be a common basket of goods. A “product” in this context corresponds to a UPC, uniquely defined by producer, weight, and size. The data were collected to help inform consumers which nearby store offered the lowest-price bundle in a given week. Consequently, stores and chains are identified by name and address, and the recorded price is the shelf price of the first unit purchased—a promotion such as “second item 50% off” is not reflected, so ICC prices largely represent non-sale prices.

The ICC sampling interval was not constant: surveyors visited every 4 weeks between January and September 2014 and every 2 weeks otherwise, so timing estimates inherit a coarser grid during most of the first post-reform year.¹⁵ Appendix A explains why collection ended in February 2015.

A second pricing dataset, from the Central Bureau of Statistics (CBS), is collected to create a truly representative basket to calculate national consumer inflation indices. It is similar to ICC but defines “products” at the product-type level—for example, a carton of omega-3-enriched brown medium eggs (any brand). This dataset forms a balanced panel and is sampled monthly, but store identities are anonymized and chain identifiers are absent. Because the CBS samples an extensive share of non-supermarkets in order to reflect a representative consumption basket, I restrict it to supermarkets and apply the same selection rules used for the ICC data. More details are provided in Appendix A.

Scanner data. StoreNext Ltd. provided daily and weekly product-store scanner extracts for 2013–2015, which I aggregate to store-weeks. After cleaning, the combined extracts contain 17 products. Store and chain identifiers cannot be matched across extracts, so the reported counts are

¹⁵Results reported in weeks that rely on ICC dates inherit this resolution; the scanner data are weekly throughout the sample.

upper bounds (Appendix A). Scanner prices equal revenue divided by units and can average over prices paid within a period (Einav et al., 2010; Strulov-Shlain, 2023). Appendix A describes the products, aggregation, cleaning, and a worked example; Section 3 reports cleaning sensitivity.

Sample construction. I construct a “Main” sample from each unbalanced source. The pricing samples retain consistently observed products and sufficiently sampled stores and carry prices backward across remaining gaps. Backward completion dates changes at the earliest period consistent with the data, so estimated adjustment speed is an upper bound. For scanner data I round to admissible denominations and retain admissible observations or sufficiently long same-price spells. Appendix A gives the exact thresholds and raw-data checks.

Table A1 in Appendix A reports the raw and main samples. The main ICC sample includes 21 products across 143 stores, the CBS sample 171 product types in 99 stores, and the main scanner sample 17 products and about 185,000 observations. Pre-reform .99 shares are 45–47% in the two pricing datasets and 23% in the scanner data; the latter records quantity-weighted transaction prices and therefore also absorbs promotions.

In some analyses of the scanner data, I restrict to a “long price spells” sample, focusing on price spells lasting at least two weeks to exclude likely promotional activity.

The ICC panel, which has chain identifiers, is the main data used for the analysis price-ending shares, the paths, and the variance decompositions; the CBS panel replicates the aggregate shares and the first-transition analysis in an independent sample; and the StoreNext scanner data, as the data source with quantities, is used for demand estimation, the demand-shortfall and detection analyses.

2.3 Left-digit bias and pricing predictions

A key feature of supermarket demand is left-digit bias: the tendency of consumers to overweight the leftmost digit of a price. Under such bias, consumers perceive prices with different leftmost digits as more different than they are in absolute terms, and prices with the same left digit as less different. This leads to downward-sloping demand curves with discontinuous drops at round-number thresholds. The implication for pricing is that firms should avoid setting prices just above those discontinuities. The literature supports this mechanism. Strulov-Shlain (2023) provides large-scale evidence from 25 U.S. supermarket chains and thousands of products supporting significant left-digit bias.¹⁶

¹⁶Early work on psychological pricing reached mixed conclusions (Basu, 1997; Gedenk and Sattler, 1999; Ginzberg, 1936) and suggested multiple potential explanations (e.g., Basu, 2006; Bizer and Schindler, 2005; Anderson and Simester, 2003). Experimental and empirical studies on 9 and 99 endings also yielded mixed evidence (Anderson and Simester, 2003; Bizer and Schindler, 2005; Sokolova et al., 2020; Thomas and Morwitz, 2005). And related work has established the presence of such bias in non-price contexts as well (Lacetera et al., 2012; Busse et al., 2013; Li and Qiu, 2023; Kraft and Rao, 2026), and extended the idea to broader demand environments (Hilger, 2018; List et al., 2023; Repetto and Solis, 2020).

To derive pricing predictions, I use the model from Strulov-Shlain (2023), where biased consumers face monopolistic firms. A price p is perceived as a weighted average of the true price and a focal reference point:

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

where $\lfloor p \rfloor$ is the floor of the price, Δ is a constant focal price ending, and θ represents the degree of left-digit bias. For instance, if $\theta = 0.2$, then a 1-cent change from 4.99 to 5.00 is perceived as a 20.8-cent jump, while a change from 5.00 to 5.01 is perceived as just 0.8 cents.

Demand is a function of perceived prices, while consumers pay the true price when purchasing. Assuming constant elasticity demand:

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

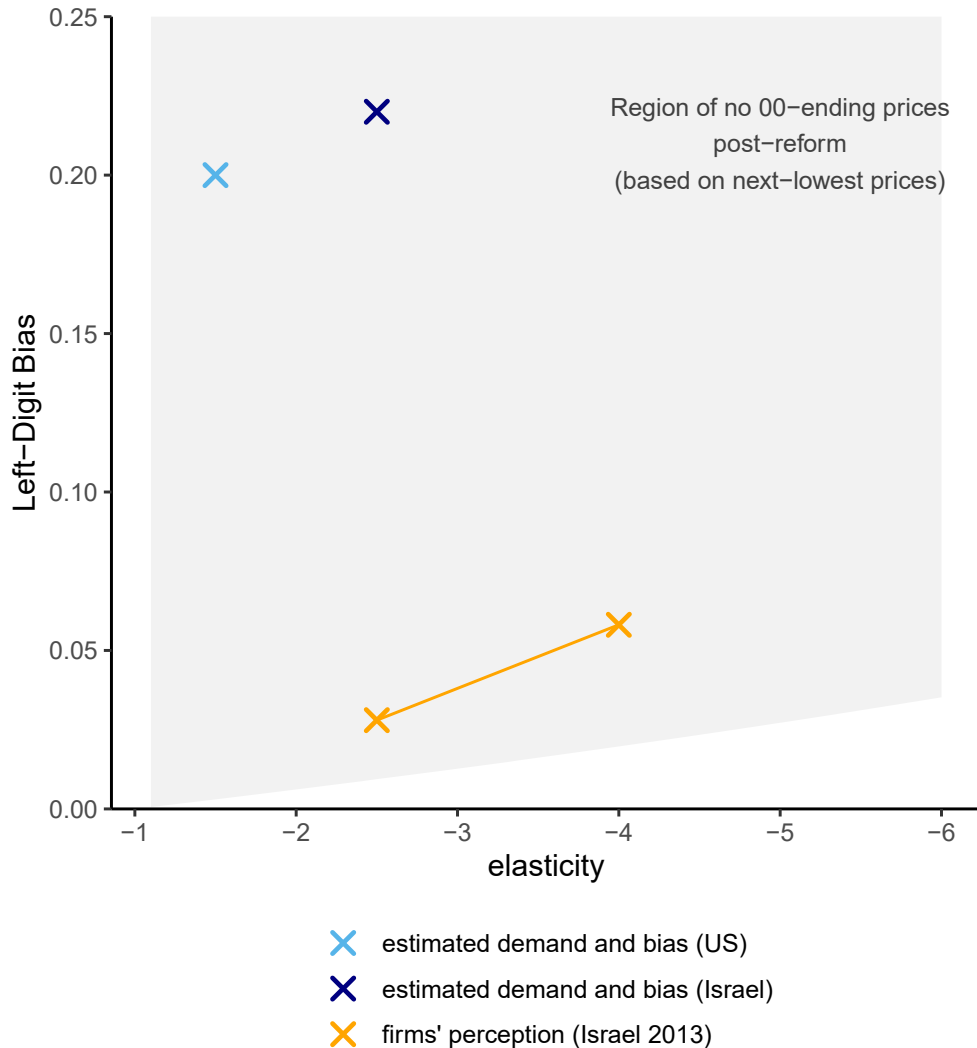
Gross profits are then $\Pi(p; \theta) = D(p; \theta)(p - c)$, assuming marginal cost c . As shown in Strulov-Shlain (2023), this model predicts bunching of prices just below round numbers—i.e., prices ending in 99 or 90—depending on available denominations.

When should .00s be avoided? A key prediction concerns the *Next-Lowest Price* P : the lowest price above a just-below price q that could be optimal for some cost level. The just-below price ends with 99 before the reform and 90 after. If P exceeds the round number (i.e., $P > \lfloor P \rfloor$), then .00 prices are never optimal—they are strictly dominated by either the just-below price or a higher price. Solving for the threshold level of left-digit bias θ_0 at which .00s become dominated yields low values. For an extreme elasticity of -7 and price level around 2, θ_0 is approximately 0.00075. For more typical elasticities of -1 to -4 , the threshold is even lower. In contrast, typical left-digit bias estimates are between 0.1 and 0.5 (List et al., 2023; Kraft and Rao, 2026; Sayag et al., 2024; Strulov-Shlain, 2023), and typical elasticities are about -1 to -4 (Butters et al., 2025; DellaVigna and Gentzkow, 2019; Hausman et al., 1994; Hitsch et al., 2021; Nevo, 2001). This calibration implies that for any reasonable levels of left-digit bias and price elasticity, *if firms optimize against the true demand function*, we should observe no .00 pricing when .99s are available.

Predictions after the reform. After the reform, the just-below price changed from .99 to .90, and the next higher admissible price from .00 to .10. These changes make just-below pricing less profitable (the gap between 4.90 and 5.00 is larger than between 4.99 and 5.00 which cuts into margins) and upward adjustments necessarily larger. Re-solving for θ_0^{Post} under these conditions raises the threshold, but it remains low: at most about 0.02 in the least favorable corner of the empirically relevant range ($\epsilon = -4$ and price level around 2), and below 0.01 at the elasticity of about -2 estimated below (Figure 1). Either way it is well below empirically observed levels of left-digit bias.

Figure 1 shows the $\theta_0(\epsilon)$ threshold values before and after the reform, alongside empirical estimates from Strulov-Shlain (2023) and the current paper (to be introduced later). The same logic that made .99s attractive before the reform also makes .90s dominate .00s after the reform.

Figure 1: White-to-Gray: Minimal levels of left-digit bias $\theta_0(\epsilon)$ above which .00 is never optimal



The figure shows the minimal levels of left-digit bias θ above which .00 prices are never optimal as the boundary between the white and gray regions. The values are a function of the price elasticity and nominal price level. The gray shading above the light line denotes the parameter set for which there should be no .00 prices after the reform.

What we learn from .00s. The prediction that .00s should be avoided provides a useful diagnostic that distinguishes the two views of pricing introduced in Section 1. If firms understand that they face left-digit biased demand—even approximately, and even if they underestimate it—they should avoid .00s both before and after the reform. Observing substantial .00 prices after the reform, when .90s are available, suggests that firms did not fully understand the demand structure, or

did not know how to optimize, or were not interested in fully optimizing. They may have known that .99s worked well without knowing *why*, and therefore could not immediately infer that .90s would dominate .00s: with .99 gone the standard of selling as the product used to sell can only be consulted *after* an ending is tried. This is the heuristic view. In contrast, the model-based view, with or without parameter uncertainty, predicts immediate shying away from .00s.

This is the sense in which the reform distinguishes heuristic pricing from model-based optimization. A rule like “use .99s” produces good outcomes before the reform but provides limited guidance afterward.

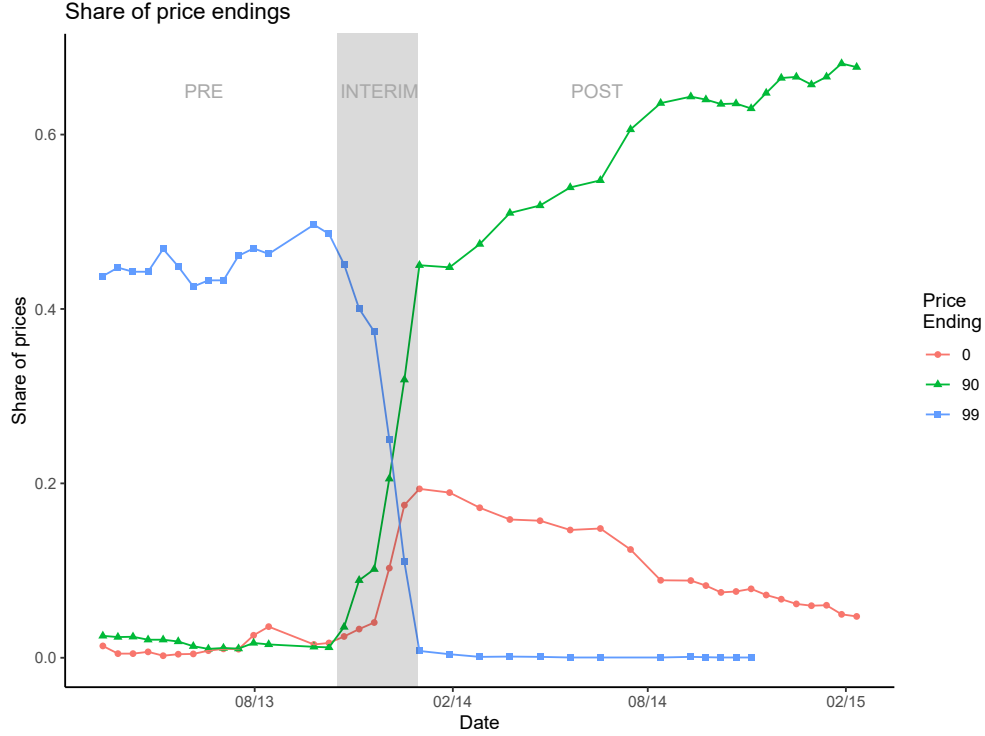
3 Results

This section presents the empirical findings on pricing patterns and demand response. I organize the results around four questions: How did price endings change after the reform? At what level are ending decisions made—product, chain, or store? What determined the initial transition to .00, and what drove the correction that followed? What was the demand consequence of choosing .00s over .90s, and did demand itself change around the reform? I close by quantifying the profit consequences, and defer interpretation to Section 4.

3.1 Price-ending shares over time

Figure 2 shows the shares of prices ending in .99, .90, and .00 by period in the ICC price data. The shaded area represents the window between the announcement (October 2013) and enactment (January 2014) of the policy change. Figure A-2 in the appendix replicates the pattern in the CBS price data.

Figure 2: Shares of price endings over time, ICC data



The figure shows the shares of .99, .90, and .00 prices across products and stores by sampling period in the ICC data. The shaded area represents the time between announcement of the reform on October 17, 2013, and its enactment on January 1, 2014. Squares represent .99 prices, triangles .90 prices, and circles are .00 prices.

Before the reform. The share of products priced at .99 was high and stable, while .00 and .90 were rare. In both pricing datasets, the share of .99 before the policy announcement was about 45% and remained stable. In contrast, the shares of .00 and .90 were low, at 1.2% and 1.7%, respectively.¹⁷ This pattern is consistent with the normative prescription from Section 2.3: under left-digit biased demand, .99 should be common and .00 rare.

Transition period. Immediately after the policy was announced, but before it was enacted, firms began changing prices. Chains had 2.5 months to update prices, and 79% of products experienced a price change in December 2013.

After the reform. When the policy took effect in January 2014, approximately 40% of prices ended in .90 and 20% in .00. The 20% share of .00s seems high: the model predicts that .00s should be avoided, yet one-fifth of prices were set at this dominated ending¹⁸. Soon thereafter, the share of .00s began a gradual and persistent decline. A year after the reform, the .00 share had fallen

¹⁷The second- and third-most frequent price endings were 49 (6.6%) and 89 (3%).

¹⁸A uniform pricing distribution implies 10%, and a uniform-if-not-90 implies less than 7%.

to about 5% in ICC data and 8% in CBS data; the share of .90s exceeded 65%, higher than the pre-reform share of .99s.

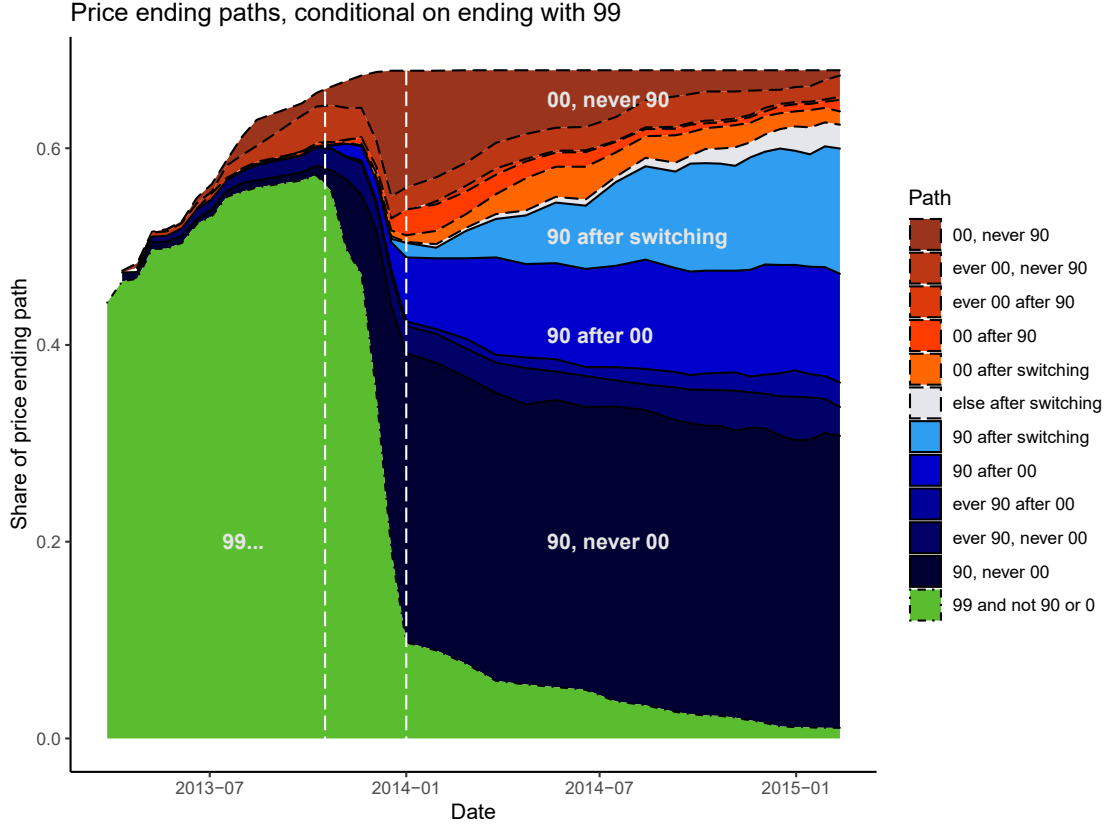
Between-chain heterogeneity. While the aggregate decline in .00s is gradual, chains exhibit more stepwise adjustments. Appendix Figure A-3 shows price-ending shares after the policy change for the six largest chains. For most chains, the .00 share drops in a stepwise fashion; for “Mega in the City” and “Rami Levy” (the second- and third-largest chains), the shares fall to nearly zero. “Shufersal” (the largest chain), with its subsidiaries “Deal” (discount stores) and “Shelli” (urban stores), reduces .00s to about 10% by mid-2014. “Wine Pavilion” and “Shufersal Deal” also appear to reduce .00s by 2015, though the data end too early to confirm this conclusively.¹⁹

3.2 Price-ending paths at the product-store level

Price-ending paths at the product-store level are informative about the adaptation process and price change decisions. Consider a product in a store initially priced at a .99. For tractability, I classify price endings into three states—90, 00, or Other—and keep track of the path each product-store follows. The state history is categorized into path histories as described by the automaton in Appendix Figure A-1.

¹⁹Appendix Figure A-5 shows qualitatively similar shares in the raw ICC data, without sample selection or imputation.

Figure 3: Price ending paths of products that ended with .99 before the policy change



The figure shows the shares of each price-ending path (described in Figure A-1). The green area is .99 or other price endings, the blue shades are .90 prices, and the red shades are .00 prices.

Figure 3 reports the shares of different price-ending paths in the ICC data. Three patterns emerge:

First, products that switch to .90 tend to stay at .90. This is evident from the large, stable dark-blue “90, never 00” segment versus the small orange “00 after 90.” Once a product is priced at a .90, it is unlikely to move to .00.

Second, products that switch to .00 tend to move to .90. The shrinking dark-red “00, never 90” segment contrasts with the growing blue “90 after 00.” Products that initially changed to .00 are 15 times more likely to end at .90 than at .00 in the last period of the data.

Third, products that switch between endings eventually settle at .90. A small set of products exhibit both transitions, switching between .90 and .00. This pattern is consistent with deliberate “RCT-like” exploration, but is rare.

3.3 Decision-making level of price endings

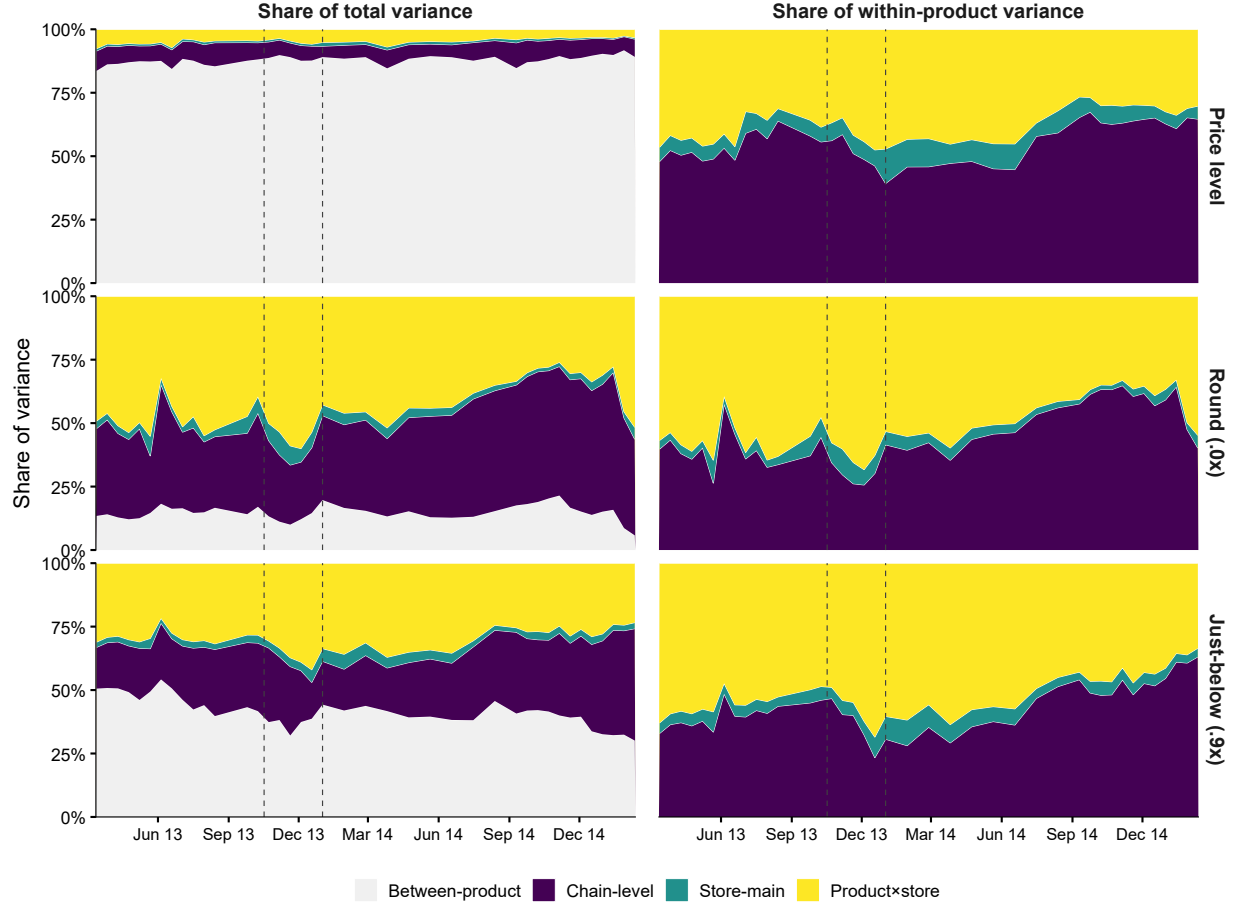
At what level are ending decisions made—product, chain, or store? Variance decompositions show that chains and products explain most of the variation in prices and ending levels, with little role for stores. The transition process into specific price endings is studied more carefully in Sections 3.4 and 3.5, which show it is organized between products rather than by chain.

In the ICC data, I decompose the cross-sectional variance of prices and of the price-*ending* choice, into nested components, separately in each biweek and restricting to the seven largest chains so that chain and store identity are distinguishable²⁰: *between-product*; product + *chain-level*, a chain’s common price or ending for a given product, shared across its stores; *+store-main*, a store’s uniform shift across its own products; and *+product*×*store*, the residual store-by-product deviation. Because products differ enormously in price level, product identity mechanically accounts for about 90% of the variance in the price *level*; the informative question is how the *within*-product variation splits between chains and stores. Figure 4 shows the decomposition for three states: the price level, the round (.0x) ending, and the just-below (.9x) ending.

Two facts emerge. First, within a product the chain is the operative unit (not the store), but there is still a lot of idiosyncrasy remaining. The chain-level component accounts for roughly one-half to two-thirds of the within-product variance in the price level and in the just-below ending, while a store’s uniform price position (store-main) is negligible throughout (under 10%). Chains price and choose endings for a given product across their own stores, and the store-level variation that does exist is product-by-store specific rather than a store-wide stance. Second, chain coordination on price endings *strengthens* as firms learn after the reform: the chain-level share of the round-ending choice rises from about two-fifths at the reform’s onset to about two-thirds by the end of 2014.

²⁰beyond the largest seven, chains only have one or two stores in the data.

Figure 4: Variance of prices and price endings, decomposed by level, biweekly



Nested decomposition of the cross-sectional variance, computed separately each biweek on the seven largest chains (so that chain and store identity are distinguishable). Rows: the price level, the round (.0x) ending, and the just-below (.9x) ending. Left: shares of total variance; right: shares of within-product variance. *Between-product* is variance across products; *chain-level* is a chain's common price/ending for a given product (across its stores); *store-main* is a store's uniform shift across its products; *product×store* is the residual store-by-product deviation. Dashed lines mark the reform announcement (Oct 2013) and enactment (Jan 2014).

Chain-level assignment leaves little store-level margin for local competition. Stores above and below the median number of competitors within 5 km differ only slightly in their initial .00 shares, and within Shufersal Deal the difference is smaller and, if anything, runs the other way (Appendix M.1, Figure A-13).²¹ Competition may still matter across chains, but the seven-chain analysis in Appendix M.1 is only suggestive.

²¹The result is qualitatively similar at other radii; a 5-km radius keeps the comparison meaningful within densely populated areas rather than separating urban from rural stores.

3.4 The transition to .00

The way the shift to .00 at the start of the reform was carried out sheds light on what was known before the reform and how learning took place. Observed covariates and additive chain or store propensities explain little of the initial transition; stable product identity accounts for most of the predictable component.

No observable characteristic predicts the initial choice of .00 or .90, and the structure that does exist is explained by product identities. The same store commonly priced one product at .90 and another at .00 in the same period. In a stacked regression across the ICC, CBS, and StoreNext samples ($N = 6,466$), the pre-reform share of .99s predicts the first-transition outcome (-0.265 , $p < 0.001$) but with product fixed effects the effect vanishes ($+0.023$, not significant); pre-reform sales, mean price, chain size, and the 5-km competitor count behave the same way. *Products* that leaned on .99 avoided .00, while within a product, stores that leaned more on .99 were no less likely to round to .00. A variance decomposition of the same outcome supports these findings: product identity explains about 75% of the explained variance out-of-sample in ICC and about half in CBS; the store share collapses from 28% to 4% under 10-fold cross-validation, so the in-sample store signal is mostly overfit; chain identity contributes about a tenth; and the observed covariates account for only 6% of the explained variance. Appendix E reports the full regressions and the decomposition.

One remaining concern is that products that changed into .90 vs .00 are different in some other ways. That is hard to rule out without random assignment. However, I check for pre-trends in demand for these subsets of products during 2013 and find no difference (Appendix D).

Systematic patterns appear in the correction (the transition out of .00), to which I turn next.

3.5 The transition out of .00

Unlike avoiding .00 in the first place, moving away from .00 does not require understanding left-digit bias, only a benchmark precise enough to detect that .00 under-performs compared to .99 while .90 does not. Pricing in the decades before the reform provided firms with an estimate of how much they sell at their familiar just-below prices, but not at .90 or .00. After the reform, a .90 with the same left digit sells about as expected, but a .00 changes the left digit and sells less than expected. I therefore test whether correction is faster where the .00 shortfall is larger and easier to detect: where the benchmark is more precise, and where the penalty is larger.

Larger shortfalls at .00 predict a faster move back to .90. Among 536 product-stores with a .00 episode (470 of which switch), a one-standard-deviation larger demand shortfall over the episode's first four weeks raises the switching hazard by 11% ($p = 0.022$), raises the probability of switching within eight weeks by 6.2 points ($p = 0.01$), and among switchers brings the switch about two weeks earlier ($p < 0.01$). The shortfall is measured relative to .90-priced peers in the same chain

$\times$ product $\times$ week cell; its construction and the full estimates are in Appendix H. Because it is measured during the episode rather than randomly assigned, I read these as predictive evidence about feedback, not as a causal effect.

Coordination takes the form of simultaneous batching rather than gradual diffusion. Decomposing each switch into cross-product spillover channels, the strongest predictor is other products switching in the same store in the same biweek (+6.3 percentage points), followed by the same store in recent biweeks (+3.9); a chain’s other stores switching contemporaneously adds +0.6, while gradual within-chain diffusion over recent biweeks is smaller and not significant (+0.15, $p = 0.17$). Appendix G reports the decomposition.

The response is concentrated where the shortfall is easier to detect. Interacting it with the precision of a product-store’s pre-reform demand (measured as the inverse dispersion of its demand residuals) gives a positive and significant interaction ($HR = 1.21$, $p = 0.001$), and the pre-reform .99-share stops predicting once precision is included. Precision is not a proxy for firm quality: if it were, precise product-stores should have avoided .00 in the first place, and the same measure does not predict who moved to .00 (Appendix H). Correction is therefore associated with detectable demand feedback rather than with knowledge of the penalty. The precision result identifies feedback-driven learning; the initial use of a dominated ending and the limited generalization across actions provide the additional separation from the structured model-based account.²²

The recovery has some regularities across products. Products that relied most on just-below pricing before the reform re-establish .90 fastest (rank correlation with the mean first-.90 date, -0.55 ; -0.80 among product-stores that eventually switch). In contrast, correcting first where .00 is dominated over the widest range of costs—the ordering a model-based firm would follow—is not supported ($\rho = +0.19$, $p = 0.40$), though with 21 products that null is imprecise. It seems that .99-reliance dominates the threshold ordering, not that the threshold ordering is rejected. Appendix I reports the timing regressions, the scanner-data version at product-specific elasticities, and the same test applied to initial adoption.

The switches were executed on chain schedules. Within a product, chain identity accounts for 44% of the variance in the date a product-store first reaches .90, against the 11% expected if stores moved on chain-independent schedules; a permutation test reshuffling chain labels within a product rejects chance at $p < 0.001$. Stores in the same chain reach .90 within a median of about nine weeks of one another, against a recovery spanning roughly a year. The same synchronization holds in the scanner data. Appendix F reports the permutation distributions.

The order in which products first adopted .00 is unrelated to the order in which they recover ($\rho = 0.07$, $p = 0.76$).

²²This is a pricing analogue of learning through noticing (Hanna et al., 2014).

3.6 Demand response to price endings

The patterns above show that firms moved away from .00s over time, and faster for products that fared relatively worse. But did .00s actually hurt demand? The model, firms’ pricing behavior before the reform, as well as prior evidence from similar settings predict they should, and yet perhaps the reform was accompanied by a change in consumer behavior (e.g., through media coverage or broader changes in the market).²³ The scanner data allow a direct test of whether .00s are associated with lower demand after the reform.

The ideal design to identify the causal effect of price endings on demand would randomize price endings.²⁴ The reform creates a quasi-experimental version of two designs. Consider a product priced at 5.99 in 2013. The first design is within-store: the same product alternating between 5.90 and 6.00 after the reform in one store. The second is an across-store design in which one store sets 6.00 and another 5.90 for the same product previously priced at 5.99 in both. Under the assumption of as-good-as random transitions into price endings (supported by Section 3.4 which showed that observed predictors explain little of the transition), the resulting within- and across-store variation is useful for estimation. It does not establish literal random assignment, so the analysis below combines fixed effects, pre-trend checks, and instruments while retaining explicit exclusion restrictions.

Using StoreNext scanner data, I estimate:

$$\log Q_{ist} = \alpha^{90} \kappa_{ist}^{90} + \alpha^{00} \kappa_{ist}^{00} + \alpha^{99} \kappa_{ist}^{99} + \beta_{ic(s)} \log P_{ist} + X'_{ist} \beta + \epsilon_{ist} \quad (3)$$

where i denotes product, s store, t week, and c chain. Q is units sold and P is the average price. For each product-store in 2013, I define the modal .99 price p_{is}^{Modal99} . The α ’s are price-ending fixed effects: $\kappa_{ist}^{99} = 1$ if, in week t , the price equals p_{is}^{Modal99} ; in 2014, $\kappa_{ist}^{90} = 1$ if the price is 9 Agorot lower than p_{is}^{Modal99} and $\kappa_{ist}^{00} = 1$ if it is 1 Agora higher. X includes controls: product-by-store fixed effects, an on-sale proxy, same-price spell-length decile-by-product fixed effects, month-by-product fixed effects, and year-by-product fixed effects.

Even with the evidence in Section 3.4, the ending dummies $\kappa_{ist}^{90}, \kappa_{ist}^{00}$ may be correlated with unobserved product-store demand: a firm might assign .00 endings to products with idiosyncratic high demand to “camouflage” the bad ending, or to products where own-demand elasticity is unusually low. To mitigate this concern, I run three instrumental-variable regressions. The first is a Hausman price IV: instrument $\log P_{ist}$ via the leave-out average price in the same chain-product-time. The second instruments κ_{ist}^{90} and κ_{ist}^{00} via the leave-out chain shares of each ending in the same product $\times$ week cell.²⁵ The third uses both IVs jointly.

²³Snir et al. (2017) study consumer perceptions to the reform, and find consistency in .90s serving as “psychological prices” instead of the .99s.

²⁴Randomization at the consumer level would also be ideal but is difficult in brick-and-mortar settings. In online marketplaces, List et al. (2023) exploit an RD design based on the smooth distribution of prices around .00s, and Dubé and Misra (2023) randomize prices in a field experiment.

²⁵If product i in store s on date t has ending .90, the leave-out instrument for κ_{ist}^{90} is the share of other stores in the same chain selling product i on date t that also use .90. This extends the leave-out-mean construction of Hausman (1996) from

A concern for the Hausman price IV is spatially correlated demand: if prices vary geographically within a chain, other stores' prices could reflect regional demand shocks rather than common costs. Two features limit this concern here. Chains price a given product similarly across their stores (Section 3.3), so the leave-out price carries chain policy rather than local conditions; and the ending IVs isolate the discrete chain-level ending component, which is further removed from a focal store's local demand. The ending instruments do not eliminate chain-wide product-specific demand shocks: identification still requires those shocks not to drive other stores' contemporaneous ending choices.²⁶

Table 1 reports the three IV strategies in three column blocks, with six specifications under the Hausman price IV (varying sample restrictions) and two specifications under each of the other two. Under the Hausman price IV alone, the demand gap between a .90 ending and a 10-Agorot-higher .00 ending for the same product ranges from +6.5% to +11.2%, with the preferred specification at +6.5% (column 5 of Table 1). Under the endings IVs alone, the gap rises to +18.3% to +22.9%, all significant at $p < 0.001$. Under both IVs together, the gap is +15.3% to +19.1%. The implied $\hat{\theta}$, computed from the most conservative set of gaps (the mapping and inputs are in Appendix C), is between 0.25 and 0.38 in five of the six Hausman columns (0.51 under the most restrictive filter)—more than an order of magnitude above the θ_0 threshold below which .00 could be optimal (at most 0.02 in the empirically relevant range; Section 2.3).

The 90-vs-00 demand gap widens when the ending choice is instrumented. This is consistent with attenuation or with firms placing .00 endings on products and weeks with relatively strong demand, although different IV specifications can also identify different weighted comparisons. I therefore treat the Hausman-only estimate as the cautious preferred estimate rather than claiming that the larger IV estimates prove attenuation.

continuous prices to the discrete ending choice. The identifying assumption is that ending choice is set centrally at the chain $\times$ product $\times$ week level; other stores' contemporaneous choices reflect chain policy and are uncorrelated with the focal store's idiosyncratic demand shock.

²⁶StoreNext store identifiers are anonymized, so a distance-based leave-out that excludes nearby same-chain stores is not feasible; the residual assumption is that chain-level cost shocks are uncorrelated with local demand shocks. Second, although mid-sample moves *into* .00 are largely store-idiosyncratic, the instrument rests on the *level* of the ending choice at a point in time, which is chain-dominated throughout (Section 3.3); the leave-out share captures that chain-level component.

Table 1: Price-ending effects — three IV strategies

	Quantity (log-units)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
α^{90}	−0.035 (0.011)	−0.014 (0.012)	−0.040 (0.026)	0.002 (0.039)	−0.036 (0.025)	−0.012 (0.025)	0.053 (0.028)	0.073 (0.029)	0.004 (0.032)	0.031 (0.033)
[p-value]	[0.001]	[0.232]	[0.116]	[0.951]	[0.147]	[0.630]	[0.057]	[0.011]	[0.898]	[0.341]
α^{00}	−0.106 (0.016)	−0.095 (0.018)	−0.122 (0.028)	−0.110 (0.032)	−0.102 (0.029)	−0.099 (0.032)	−0.130 (0.036)	−0.156 (0.036)	−0.149 (0.035)	−0.160 (0.035)
[p-value]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[0.002]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
α^{99}	−0.074 (0.018)	−0.064 (0.018)	−0.067 (0.014)	−0.102 (0.024)	−0.070 (0.015)	−0.064 (0.016)	−0.074 (0.013)	−0.067 (0.014)	−0.068 (0.015)	−0.060 (0.016)
[p-value]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
90 vs 00	0.070 (0.018)	0.081 (0.019)	0.081 (0.031)	0.112 (0.036)	0.065 (0.031)	0.087 (0.031)	0.183 (0.040)	0.229 (0.042)	0.153 (0.041)	0.191 (0.040)
[p-value]	[<0.001]	[<0.001]	[0.008]	[0.002]	[0.038]	[0.005]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
99 vs 00	0.031 (0.024)	0.030 (0.025)	0.054 (0.030)	0.008 (0.033)	0.031 (0.029)	0.035 (0.030)	0.056 (0.038)	0.089 (0.037)	0.081 (0.037)	0.100 (0.037)
[p-value]	[0.187]	[0.223]	[0.068]	[0.804]	[0.282]	[0.242]	[0.141]	[0.017]	[0.030]	[0.007]
Hausman IV (chain × price)	X	X	X	X	X	X			X	X
Endings IV							X	X	X	X
Spell-length filter		≥2				≥2		≥2		≥2
99 pre prevalence			≥1	≥5	≥5	≥5	≥5	≥5	≥5	≥5
90/00 post prevalence				≥5 both	≥5 either	≥5 either	≥5 either	≥5 either	≥5 either	≥5 either
First-stage F (min)	—	—	—	—	—	—	59,682	49,698	—	—
N	247,003	207,935	247,003	247,003	247,003	207,935	246,078	206,087	246,078	206,087

The table shows the demand specification in Equation 3 under three IV strategies, with cluster-robust standard errors

clustered by store in parentheses and exact p -values in brackets. Columns 1–6 use the chain-interacted Hausman price IV under six prevalence filters. Columns 7–8 use the pooled-price IV plus leave-out chain ending IVs for κ^{90} and κ^{00} . Columns 9–10 use both: chain-interacted price IV jointly with the ending IVs. The first-stage F (minimum across the three endogenous regressors) is reported for the pooled-price ending-IV columns; the leave-out chain-share instruments are strong because chains set endings centrally. The chain-interacted specifications (columns 1–6 and 9–10) involve over a hundred simultaneous first stages, for which no single summary statistic is standard; per-equation first-stage statistics are available in the replication package. The 90-vs-00 gap is the left-digit-bias measure and is larger when the ending choice is instrumented.

The *levels* of the three ending coefficients are mostly negative, which reflects the comparison group rather than the endings themselves. The omitted weeks are those at other prices within the product-store; these include promotional weeks whose demand lift exceeds what the price elasticity alone predicts, so weeks at the regular, modal-based prices look low by comparison. This contamination is common to all three ending dummies and differences out of the pairwise contrasts, which is why the 90-vs-00 and 99-vs-00 gaps are the objects of interest. Under the ending IVs α^{90} turns positive, consistent with the direction of selection noted above: uninstrumented .90 weeks understate demand at .90.

Varying data cleaning choices is consistent with attenuation from short-spell and ambiguous-price observations. The cleaning rule drops a meaningful share of observations; re-running the chain-interacted Hausman IV specification at five thresholds gives a monotone pattern, from +9.5%

under the strictest rule (27% of observations kept) to +2.1% with no filter at all (96% kept), with the default rule in between. Appendix O reports all five with standard errors.²⁷

3.7 Was demand stable around the reform?

A reasonable alternative explanation is that consumer attention to right digits temporarily shifted with the reform’s media coverage, which would contaminate the estimated penalty. I approach this in three ways.

First, and most simply, I estimate the penalty period by period. As time moves further from the reform, the observations at .00 become more and more selected (as shown above), so this is imperfect. Still, splitting the post-reform period into 2014H1, 2014H2, and 2015 and re-estimating gap-by-period under the endings-IV specification gives +0.198 ($p < 0.001$) in 2014H1, +0.161 ($p = 0.016$) in 2014H2, and -0.025 ($p = 0.83$) in 2015. The stability claim rests on the two precisely estimated cells: the 2014H1 and 2014H2 gaps do not differ ($p = 0.64$), so there is no evidence for the gap getting stronger as the reform’s news cycle recedes. The 2015 cell is uninformative about the structural penalty: by then few product-stores still use .00, and those that remain are a selected set, the product-stores whose .00 endings were not underperforming.

Second, I train a demand model on 2013 only (log Q on log p , a sale flag, month-of-year and Jewish-holiday controls, and product $\times$ store fixed effects, with no ending terms) and project it forward (details in Appendix J.3). Pooled monthly residuals average +0.01 on the 2013 hold-out and drift to -0.05 by 2015. This modest drift is consistent with broad economic stability, though it does not by itself prove that every demand parameter was unchanged. Split by ending class, prices that moved to .00 sold 0.14 log points below the model’s prediction in the first half of 2014—in the range of the ending-IV estimates of Table 1—and the difference disappears over 2014–15 as underperforming .00s are abandoned. Survivorship provides a plausible explanation for the disappearance because firms retained .00 where their own data showed less shortfall; the residual exercise cannot separately identify survivorship from a change in the penalty.

Third, the price elasticity itself is likewise stable. Estimating it separately by period (2013, 2014H1, 2014H2, 2015) under the period-interacted Hausman IV gives -2.01 , -2.00 , -2.01 , and -2.02 : the largest gap is 0.03, about 1% of the level. A formal test rejects exact equality, but a difference of 0.03 in an elasticity of -2 has no economic content: neither the θ_0 thresholds nor the profit calculations move over that range (Appendix J).

Taken together, these tests provide no evidence that a temporary attention shift reversed the .90-versus-.00 ranking, which is what the qualitative results require. They cannot exclude every change in consumer attention, but the stability of the elasticity, the substantial and similar .90-versus-.00 gaps in the two precisely estimated periods, and the modest out-of-sample drift make

²⁷These shares use the StoreNext panel before the additional demand-regression filters; Table A1 reports a narrower descriptive panel that also requires consistent product coverage, which explains its smaller StoreNext count.

a ranking reversal less plausible than economic stability. Interestingly, the press treated round-ups as the objectionable move, presenting a change from 5.99 to 6.00 as firms using the reform to raise prices (footnote 14). Publicity of that kind would make consumers more sensitive to .00 rather than less, which is the opposite of what the attention alternative requires. The estimated gap is, if anything, larger in the half-year when coverage was heaviest (+0.198 in 2014H1 against +0.161 in 2014H2), though the two do not differ statistically ($p = 0.64$). Because the penalty is estimated from transition-period variation, a temporarily lower effective bias would move the estimated penalty and the computed transition losses together; under that account, both would understate the corresponding steady-state penalty.^{28,29}

3.8 Quantifying forgone profits

I quantify the profit consequences of these patterns by comparing observed pricing to a counterfactual of full optimization.

Without cost data I cannot do the comparison using the data directly, as I do not observe markups. Instead, I use a two-product logit demand model with left-digit bias, calibrated to match the estimated price elasticity and left-digit bias from the data. For each of many simulated cost draws, I compute the profit-maximizing price and compare it to the price implied by a descriptive “rule-of-thumb” heuristic that matches observed pricing patterns. The heuristic extends Strulov-Shlain (2023): firms compute a naive optimal price ignoring left-digit bias, then round to .99/.90s or .00s with probabilities estimated month-by-month from the data. Forgone gross profits equal the percentage gap between profits under heuristic pricing and profits under optimal pricing. The forgone profits are calculated relative to the best the firm can do in each month: before the reform firms can price to the cent, afterwards only to the dime. Therefore the calculated gap reflects purely the satisficing penalty. A complete description of the demand model, calibration, heuristic specification, and computation appears in Appendix P; robustness analyses augmenting it appear in Appendix N. This is a conservative exercise: it yields smaller losses than Strulov-Shlain (2023).

²⁸A related channel is payment method: cash transactions were already rounded to the nearest 10 Agorot at the register before the reform (Section 2), so the reform changed posted prices rather than the amounts cash buyers paid. Any payment-mix channel would therefore have to operate through the perception of posted prices, which is what the ending coefficients capture.

²⁹Firms may choose .00 endings to placate consumers during a confusing policy change. Section 4 develops this point and the related price-image alternative.

Figure 5: Estimated counterfactual forgone gross profits month by month

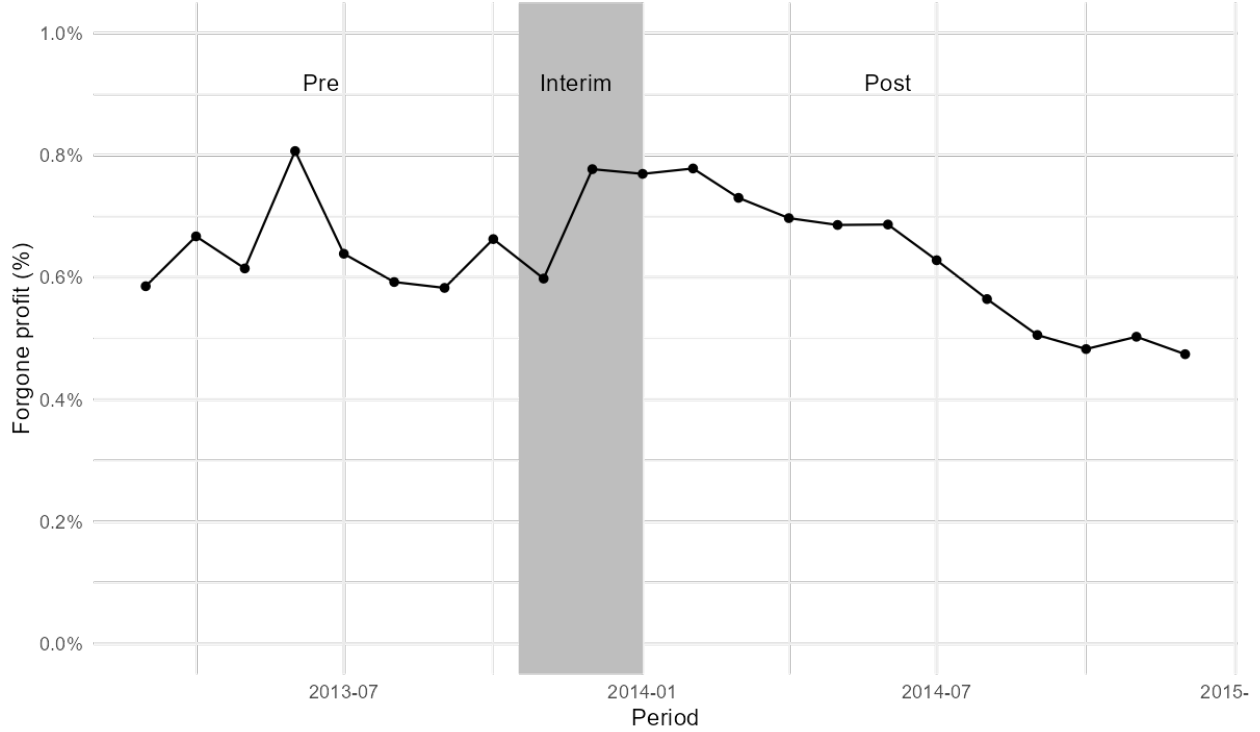


Figure shows counterfactual profits under a monthly-estimated predictive rule-of-thumb pricing relative to full optimization, with the own-price coefficient set so the simulated mean elasticity equals the estimated -2.0 and forgone profit averaged across the cost draws. Pre-reform forgone profits hover around 0.6% ; rise to about 0.8% at the reform's onset; decline to about 0.5% by the end of 2014, below the pre-reform baseline.

Figure 5 shows the results. Before the reform, forgone profits hover around 0.6% —firms were not at the optimum, but not far from it.³⁰ At the reform's onset, forgone profits rise to about 0.8% as firms initially chose dominated .00s. Over the following year, forgone profits gradually decline as firms learn, eventually falling to roughly 0.5% by the end of 2014—*below* the pre-reform level.³¹ The reform's forced exploration moved firms past the pre-reform local optimum and toward the global one, with the remaining gap potentially harder for firms to detect.

The level of forgone profit is sensitive to the demand parameters, but the qualitative pattern is stable across the specifications examined. Across a 3×4 grid crossing $\theta \in \{0.11, 0.22, 0.36\}$ with a mean elasticity of $\{-1.5, -2.0, -2.5, -4.0\}$, the post-reform decline below the pre-reform level occurs in every cell, under both the regime-constrained benchmark and the always-cents-unconstrained benchmark (Figure A-14 in the appendix). The headline Figure 5 is the $(\theta = 0.22, \varepsilon = -2.0)$ cell of this grid; $\theta = 0.22$ sits just below the estimated range of 0.25 – 0.38 . Forgone profit

³⁰Whether a loss of this size is “large” or “small” depends on context. It is smaller than many estimates of pricing inefficiencies in other settings, but represents meaningful forgone revenue for a competitive industry with thin margins: 0.5% of gross profits can be 5% or more of net profit.

³¹A second benchmark, the always-cents-optimal counterfactual, places the cost of the reform constraint alone at about 0.15 percentage points throughout the post-reform period, on top of the heuristic cost (Figure A-15a).

rises with $|\varepsilon|$ —the -4.0 cell is several times the -2.0 level—and under constant-elasticity (log-log) demand the levels are larger still with the same time-series pattern. Smoothing the period-by-period heuristic parameters produces a similar post-reform path (Appendix N). These are rough model-based estimates rather than precise measurements but they give a reasonable magnitude of effects. Further, the robust result within the exercises is the time-series pattern: a rise at the reform followed by improvement below the pre-reform benchmark.

4 Interpretation

Section 3 established multiple facts that an account of firm pricing must jointly explain. Firms priced at .99 for decades and almost never at .00. When .99 was banned, about a fifth of prices moved to the dominated .00 rather than to .90, and which ones did has little structure in observed covariates or in additive chain and store propensities, though product identity accounts for most of the predictable component. Those .00 prices then reverted to .90 over roughly a year, at a pace predicted by the demand shortfall each firm could detect against its own past sales, and fastest where that benchmark was most precise. The estimated demand relationship was economically stable across the reform, and the preferred specification puts .00 at least 6.5% below .90 in the periods that identify it. And the episode left firms better off than before, with estimated forgone profits ending below their pre-reform level.

The biggest puzzle is the move to .00s. Within the demand class of Section 2.3, a firm that understood demand would have recognized immediately that .90 dominates .00—the same logic that made .99 attractive before the reform. The reason is that pricing under left-digit bias has discontinuities at round-number thresholds, so the usual intuitions that optimal prices move smoothly with costs and that small price changes have small effects do not apply. This makes this a hard problem for firms and provides identification for the researcher. Firms were able to solve this hard problem by rule of thumb before the reform, yet failed to extend the solution afterwards.

What kind of pricing process produces an idiosyncratic move to a dominated ending, corrected product by product in a way predicted by detectable shortfalls? Table A4 lays out the candidate accounts and the evidence that distinguishes them. The evidence cannot logically exclude every flexible alternative, but that does not put the accounts on equal footing: several fit important features poorly, while others fit only by adding enough unobserved structure to lose empirical bite (Appendix L). Section 4.1 states the two main views and what each predicts; Sections 4.2–4.4 take up the three alternatives that come closest—optimization under uncertain or unrestricted beliefs, deliberate experimentation, and decision costs.

Table A4 in Appendix L lays out thirteen accounts of the transition to .00, stating for each what it predicts, what the data show, and where the paper treats it. Twelve are alternatives to the paper’s own account: full-information optimization; stable but wrong parameter beliefs; structured

optimization under parameter uncertainty; non-monotonic demand with ending-specific intercepts; chain-wide belief heterogeneity; deliberate experimentation; a wholesale-cost shift; manufacturer-dictated endings; consumer placation or price image; confusion about competitors rather than demand; decision or attention costs; and random implementation errors. The five eliminated without discussion below are worked through in that appendix.

4.1 Heuristics versus model-based optimization

Section 1 defined the two views. Both produce .99 and no .00 in steady state, so pre-reform behavior is consistent with either, but they separate once .99 is banned, because they differ in whether a firm can rank an ending it has never used.

The version of model-based optimization tested here is a belief-based max operator whose demand model slopes downward with possible discontinuities at round numbers, and whose uncertainty concerns parameters but not the model itself.³² The reform rejects the first of its two predictions and provides weaker evidence against the second. Parameter beliefs in the range consistent with pre-reform .99-use imply avoiding .00 from the start; instead a fifth of prices moved there. And because the model links all prices, correction should proceed in the order of expected forgone profit and re-rank related products together; instead it followed prior .99-reliance more closely than the threshold ordering, and proceeded product by product, although the 21-product test is too imprecise to reject that ordering on its own (Section 3.5). Allowing frictions does not repair either prediction: a firm that knows demand but cannot afford to optimize continuously settles near the model’s prescription, not at a dominated ending.

Heuristic pricing is a parsimonious explanation of what happened. A rule that clears its benchmark persists without the firm understanding *why* it works. At the same time, a rule that falls short triggers local search among nearby alternatives.³³ With .99 banned, no benchmark exists for either replacement until one is tried, so the initial choice carries no information about relative value, which is why observed covariates predict it so poorly. Once tried, the firm is able to use its experience to recognize that .00 sells below what the product used to sell, and changes the ending. Note that nothing in this account requires the firm to know in advance that .90 dominates .00.

A natural objection is that firms care about performance and would surely notice a mistake this large. They do care, but they can notice mistakes only through comparisons they can actually run, and they do not know what the maximum is. Such a comparison says nothing about .90 versus

³²An unrestricted model—e.g., one with a free intercept for every ending—can rationalize any behavior and therefore predicts nothing here. This is further developed in Section 4.2.

³³This description aligns with model-free reinforcement learning (Sutton and Barto, 2018) with an endogenous aspiration level (Börgers and Sarin, 2000): actions are reinforced or abandoned depending on whether realized payoffs clear a standard that itself adapts to experience. The agent does not use a model to generalize across actions. Nonetheless, under most reinforcement learning procedures, dominated actions are not sampled.

.00 before either is tried. Detection is accordingly stronger where the benchmark is more precise, which is what the reversion patterns show (Section 3.5).³⁴

The reform’s evidence also lines up with the comparative statics of satisficing when alternatives differ in how noisily their values are assessed. In the framework of Salant and Spenkuch (2026), each alternative carries a value and an evaluation noise that grows with the object’s complexity, and a satisficer accepts the first alternative whose perceived value clears an aspiration; they test it among chess players, where complexity varies across moves while choice sets stay fixed. The reform maps onto similar primitives, with unfamiliarity (more than complexity) concentrated in the untried option. A .90 keeps the left digit, so its value follows from decades at .99 by a small extrapolation; a .00 crosses the threshold, so its value turns on the size of a discontinuity the firm rarely observed. With the aspiration anchored on the product’s own history, the framework delivers the reform’s patterns: take-up of the dominated ending governed by how hard it is to evaluate rather than by the profit gap (initial adoption is unrelated to the threshold ordering; Section 3.5), correction through a local audit of the action taken, paced by the precision of the firm’s own benchmark (Table A3), and search that stops once .90 clears the benchmark, with little further alternation between endings (Section 3). The identification is complementary: Salant and Spenkuch (2026) vary evaluation complexity holding the choice set fixed, while the reform removes the satisfactory action holding demand fixed—a margin on which the aspiration itself must adjust, and which a fixed-aspiration design does not reach.

4.2 Could wrong beliefs explain the patterns?

One might wish to preserve the model-based framework by allowing firms to hold incorrect beliefs. Perhaps firms underestimated left-digit bias enough, leading some to believe .00s were acceptable. However, using the Method of Simulated Moments as in Strulov-Shlain (2023) (details in Appendix Q), I estimate that firms priced as if they believed $\theta = 0.02$ – 0.04 , much smaller than the demand-side estimates of $\theta = 0.25$ – 0.38 , but still positive. Within the specified demand class, even these underestimated beliefs are inconsistent with .00s post-reform: the perceived-bias estimates sit inside the no-.00 region of Figure 1 at every elasticity (with elasticity of -2 the threshold is below 0.01). If firms carried those stable parameter beliefs across the reform and optimized against them, they would have avoided .00s after the reform just as they did before.

In addition, there is no sufficient chain-wide heterogeneity. All chains used many .99s before, all had excess .00s after the reform, and additive chain identity explains little of which products initially received .00. In other words, there is no evidence of some chains “getting it right” and others acting randomly. Data suggest that firms held *precise* beliefs about their familiar price and only

³⁴That the correction is paced by benchmark precision separates feedback-driven accounts from feedback-free ones, but it does not separate heuristic pricing from learning models generally, since any learner extracting a signal from noisy data responds faster where the data are less noisy (see the note to Table A3). The precision result establishes that firms were reacting to realized demand at all.

diffuse beliefs about the untried .00; the reform’s abundant .00 pricing sharpened the latter, which is what predicts how fast each detected the .00 mistake and pushed firms closer to the optimum.

A separate issue is that an excess of *both* .00- and .90s immediately after the reform is difficult to reconcile with a single stable left-digit-bias parameter in this model: a belief low enough to permit the .00 mass implies far less bunching at .90 than observed, and a belief high enough to produce the .90 bunching implies avoiding .00 entirely. The heuristic interpretation that firms knew .99s worked without knowing why provides a more parsimonious account of both patterns.

In fairness, a flexible enough model-based alternative can always explain the patterns. Suppose, for example, subjective demand is non-monotonic and gives each price ending its own demand intercept. A firm can then believe that .00 is best initially and revise that intercept downward after observing poor sales, reproducing the product-by-product return to .90. The data cannot disprove this account. I nevertheless view it as less plausible than the heuristic interpretation. It must assign sufficiently favorable latent values to .00 despite decades of successful .99-pricing and then allow those values to vary and update at the product or product-store level to match the idiosyncratic onset and correction. More fundamentally, once each ending has an essentially unrestricted value that is learned only when that action is tried, the model supplies little *ex ante* discipline over choices and becomes indistinguishable from a model-free action-value learning. It shows that an optimizing representation can be written down *ex post* without offering restrictions that give a model-based explanation empirical bite.³⁵

4.3 Could the .00 mass be deliberate experimentation?

Firms may have believed the reform changed demand and assigned .00 deliberately to learn. The observed assignment and updating are difficult to reconcile with a centrally designed experiment intended to learn quickly: there are few repeated .90/.00 comparisons, product identity organizes assignments across competing chains, available cross-store contrasts do not produce rapid chain-wide updating, and testing usually stops once .90 is tried. This remains compatible with experimentation under slow information aggregation. Likewise, a chain-wide headquarters belief shock combined with centralized experimental assignment would predict a stronger additive chain component than the data show (Appendix L).

Adaptive bandit learning fits the trial and error (Rothschild, 1974; den Boer, 2015), but a classic independent-arm bandit is a maximizing procedure on the unstructured side of this paper’s distinction: evidence about one ending does not re-rank another. Structured pricing models instead link candidate prices, and Thompson sampling assigns substantial probability to .00 only if firms believe it may have high value (Harrison et al., 2012; Russo and Van Roy, 2014; Schulz et al.,

³⁵The particular misspecification in which consumers favor round prices is less flexible and fits the evidence poorly: decades of pre-reform .99-pricing and the negative estimated response to .00 weigh against it.

2020; Misra et al., 2019). The evidence therefore favors relatively unstructured learning over reoptimization within the maintained structured-demand class.

4.4 Could decision costs rationalize choices?

Another alternative accepts the behavior but argues that if attention is scarce and pricing is complicated, then following rules of thumb and settling for good enough may itself be the optimal procedure, and the observed losses are the price of economizing on decisions rather than evidence of missing knowledge.³⁶ This is the rational-inattention view (Sims, 2003; Maćkowiak and Wiederholt, 2009; Matějka, 2015). It is worth separating two versions of this view. One asks whether decision costs can rationalize the specific choice of .00 over .90; the evidence says no. The other asks whether the firm’s broader policy—price by rule of thumb rather than estimate demand—is optimal given what learning demand actually costs. That question requires knowing costs I do not observe, and the paper does not answer it.

Observed implementation costs do not favor .00 over .90. The reform required chains to reprice almost every product, and setting .90 appears no more costly to implement than setting .00. Menu costs and coarse adjustment therefore do not by themselves explain selection of .00 under the maintained payoff ranking. The product-specific initial choices, response to short-run demand shortfalls, benchmark-precision result, and discrete chain switches also weigh against a single attention bottleneck, while leaving product-specific evaluation costs possible (Appendix L).

The No Improving Action Switches condition of Caplin and Dean (2015) gives the conditional formal point: under the payoffs implied by the structured subjective-demand class and stable beliefs across the reform, switching to .00 reduced profits, so a costly-information model built on those payoffs does not select it. Models with different or reform-shifted subjective payoffs remain available and inherit the flexible-belief requirements of Section 4.2.

Thus, the remaining explanation is that learning demand is at least *perceived* to be costly. Whether declining to learn it was the right call depends on costs that are not observable here, and the paper deliberately takes no position. The claim is that a model of demand with enough structure and consistent with the data was not learned. The claim is not that it should have been learned.

5 Discussion

This paper studies how supermarket chains adjusted prices following a reform that restricted available price endings. The evidence suggests firms arrived at their pricing strategies through adjusting heuristics rather than by learning demand and optimizing. When the reform removed

³⁶This route also faces a regress: if deciding is costly, deciding how much to decide is costly too (Conlisk, 1996).

.99s, many firms initially chose dominated .00s before gradually learning to shy away from them. The adjustment took about a year and involved meaningful profit losses.

5.1 Implications for research

Pre-reform bunching at .99 and the calibrated forgone-profit gap of about 0.6% make optimization against left-digit-biased demand a useful approximation to steady-state pricing. The same model does not reproduce the full price distribution without perceived bias differing from estimated consumer bias, and stable parameter beliefs within it do not predict the transition to .00. Optimization-based predictions may therefore be less reliable during transitions if firms adjust rules without re-estimating demand. Counterfactuals for policy changes, entry, mergers, and demand shocks should allow conduct and learning to change with the feasible optimum (Bhattacharya et al., 2026).

5.2 Implications for practice

Heuristic pricing can work well in stable environments but adapt slowly after a choice-set change. The mechanism implies that a demand representation that ranks newly available actions could shorten adjustment, while deliberate experimentation may reduce losses by improving the benchmarks used to evaluate current rules (Hanna et al., 2014; Schwartzstein, 2014; Gagnon-Bartsch et al., 2023; Dubé and Misra, 2023; List et al., 2023). In this episode, reform-induced exploration was beneficial under the maintained demand and cost assumptions: estimated forgone profits ended below their pre-reform level, about 0.5% versus 0.6%.

5.3 Limitations

Several limitations apply. First, this is a single case study in one industry and country. The specific patterns depend on the particular reform and demand structure. Whether similar dynamics occur in other settings is an empirical question.

Second, the magnitudes depend on assumptions. The forgone profit estimates rely on structural assumptions about demand and costs. The qualitative pattern of baseline deviation, spike at reform, gradual improvement is more robust than specific magnitudes. Relatedly, the demand estimates rest on 17 products: they support a pooled penalty estimate, not product-level generalization to a supermarket’s full assortment (Appendix K).

Third, as with most observational evidence, the paper cannot logically exclude every alternative. That limitation does not make the alternatives equally probable. Section 4 and Table A4 show that the natural alternatives either fit the combined evidence less well or require additional, largely unrestricted beliefs to explain the initial .00 mass, the eventual return to .90s, and the product-store patterns. For that reason, I view benchmark-audited rules of thumb as the most plausible account,

while acknowledging that the NIAS and dominance arguments are conditional on the structured demand class and stable subjective payoffs and that the IV estimates retain exclusion restrictions for chain-wide product-specific shocks.

Finally, some may object that price endings are too minor to matter, making this case unrepresentative of firm decision-making more broadly. The estimated gap shows that endings matter for the affected scanner-sample products, and related evidence finds they matter in other settings (List et al., 2023; Strulov-Shlain, 2023). Whether firms treat similarly consequential margins in the same way remains an external-validity question.

5.4 Conclusion

Surveys and interviews of businesspeople support heuristic-based pricing (Bewley, 2025; Haynes, 1964). This paper identifies a specific learning mechanism: firms can use local experience to find pricing rules that work without learning the demand structure that makes them work.

The reform exposed the limit of that process. Firms had to adjust when .99 became unavailable, and the resulting price distribution generated estimated forgone profits below the pre-reform level by the end of 2014. That comparison is conditional on the maintained demand, cost, and heuristic specifications. Observing near-optimal behavior therefore does not imply knowledge of the optimum: trial and error can produce reasonable strategies while leaving firms vulnerable to change and short of the frictionless benchmark when exploration stops too soon.

References

- Almunia, Miguel, Jonas Hjort, Justine Knebelmann, and Lin Tian**, “Strategic or Confused Firms? Evidence from “Missing” Transactions in Uganda,” *Review of Economics and Statistics*, 2024, *106* (1), 256–265.
- Anderson, Eric T. and Duncan I. Simester**, “Effects of \$9 Price Endings on Retail Sales: Evidence from Field Experiments,” *Quantitative Marketing and Economics*, 2003, *1* (1), 93–110.
- Aparicio, Diego and Roberto Rigobon**, “Quantum Prices,” *Journal of International Economics*, 2023, *143*, 103770.
- Artinger, Florian M. and Gerd Gigerenzer**, “Heuristic Pricing in an Uncertain Market: Ecological and Constructivist Rationality,” 2016. Working paper.
- , —, and **Perke Jacobs**, “Satisficing: Integrating Two Traditions,” *Journal of Economic Literature*, 2022, *60* (2), 598–635.
- Ater, I and O Gerlitz**, “Round prices and price rigidity: Evidence from outlawing odd prices,” *J. Econ. Behav. Organ.*, 2017.
- Ater, Itai and Oren Rigbi**, “Price Transparency, Media and Informative Advertising,” 2019.
- Basu, Kaushik**, “Why are so many goods priced to end in nine? And why this practice hurts the producers,” *Economics Letters*, 1997, *54* (1), 41–44.
- , “Consumer Cognition and Pricing in the Nines in Oligopolistic Markets,” *Journal of Economics & Management Strategy*, 2006, *15* (1), 125–141.
- Bewley, Truman F**, *Price setting*, Oxford, England: Polity Press, May 2025.
- Bhattacharya, V, G Illanes, A A Kreps, J Salas, and D Stillerman**, “A Large-Scale Evaluation of Merger Simulations,” *Available at SSRN 5177604*, 2026.
- Bizer, George Y. and Robert M. Schindler**, “Direct Evidence of Ending-Digit Drop-Off in Price Information Processing,” *Psychology and Marketing*, 2005, *22* (10), 771–783.
- Blettner, Daniela P., Zi-Lin He, Songcui Hu, and Richard A. Bettis**, “Adaptive Aspirations and Performance Heterogeneity: Attention Allocation among Multiple Reference Points,” *Strategic Management Journal*, 2015, *36* (7), 987–1005.
- Bloom, N and J Van Reenen**, “Measuring and explaining management practices across firms and countries,” *Q. J. Econ.*, 2007.
- Börgers, Tilman and Rajiv Sarin**, “Naive reinforcement learning with endogenous aspirations,” *Int. Econ. Rev.*, November 2000, *41* (4), 921–950.
- Busse, Meghan R, Nicola Lacetera, Devin G Pope, Jorge Silva-Risso, and Justin R Sydnor**, “Estimating the effect of salience in wholesale and retail car markets,” *American Economic Review*, 2013, *103* (3), 575–79.
- Butters, R Andrew, Daniel W Sacks, and Boyoung Seo**, “Why Do Retail Prices Fall During

- Seasonal Demand Peaks?," *The RAND Journal of Economics*, 2025, 56 (1), 35–54.
- Caplin, Andrew and Mark Dean**, "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 2015, 105 (7), 2183–2203.
- Cho, Sungjin and John Rust**, "The flat rental puzzle," *The Review of Economic Studies*, 2010, 77 (2), 560–594.
- Conlisk, John**, "Why Bounded Rationality?," *Journal of Economic Literature*, 1996, 34 (2), 669–700.
- Cyert, Richard M. and James G. March**, *A Behavioral Theory of the Firm*, Englewood Cliffs, NJ: Prentice-Hall, 1963.
- DellaVigna, Stefano and Matthew Gentzkow**, "Uniform Pricing in US Retail Chains," *The Quarterly Journal of Economics*, 06 2019.
- den Boer, Arnoud V.**, "Dynamic Pricing and Learning: Historical Origins, Current Research, and New Directions," *Surveys in Operations Research and Management Science*, 2015, 20 (1), 1–18.
- Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes**, "Just Starting Out: Learning and Equilibrium in a New Market," *Am. Econ. Rev.*, 2018, 108 (3), 565–615.
- Dube, Arindrajit, Alan Manning, and Suresh Naidu**, "Monopsony and Employer Mis-optimization Account for Round Number Bunching in the Wage Distribution," *Unpublished manuscript*, 2017.
- Dubé, Jean-Pierre and Sanjog Misra**, "Personalized Pricing and Consumer Welfare," *J. Polit. Econ.*, January 2023, 131 (1), 131–189.
- Einav, Liran, Ephraim Leibtag, and Aviv Nevo**, "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?," *QME*, 2010, 8 (2), 207–239.
- Gabaix, Xavier**, "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 2014, 129 (4), 1661–1710.
- Gagnon-Bartsch, T, M Rabin, and J Schwartzstein**, "Channeled attention and stable errors," *mimeo*, 2023.
- Gedenk, Karen and Henrik Sattler**, "The impact of price thresholds on profit contribution should retailers set 9-ending prices?," *Journal of Retailing*, 1999, 75 (1), 33–57.
- Gertler, Paul, Sean Higgins, Ulrike Malmendier, and Waldo Ojeda**, "Do Behavioral Frictions Prevent Firms from Adopting Profitable Opportunities?," Working Paper 33387, National Bureau of Economic Research 2025.
- Ginzberg, Eli**, "Customary prices," *The American Economic Review*, 1936, 26 (2), 296–296.
- Goldfarb, Avi and Mo Xiao**, "Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets," *American Economic Review*, 2011, 101 (7), 3130–61.

- Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein**, “Learning through noticing: Theory and evidence from a field experiment,” *The Quarterly Journal of Economics*, 2014, *129* (3), 1311–1353.
- Harrison, J. Michael, N. Bora Keskin, and Assaf Zeevi**, “Bayesian Dynamic Pricing Policies: Learning and Earning Under a Binary Prior Distribution,” *Management Science*, 2012, *58* (3), 570–586.
- Hausman, Jerry A**, “Valuation of new goods under perfect and imperfect competition,” in “The economics of new goods,” University of Chicago Press, 1996, pp. 207–248.
- Hausman, Jerry, Gregory Leonard, and J Douglas Zona**, “Competitive analysis with differentiated products,” *Annales d’Economie et de Statistique*, 1994, pp. 159–180.
- Haynes, W Warren**, “Pricing practices in small firms,” *Southern economic journal*, April 1964, *30* (4), 315.
- Hendel, Igal, Saul Lach, and Yossi Spiegel**, “Consumers’ activism: the cottage cheese boycott,” *The RAND Journal of Economics*, 2017, *48* (4), 972–1003.
- Hilger, Nathaniel**, “Heuristic Thinking in the Market for Online Subscriptions,” *Available at SSRN 3296698*, 2018.
- Hitsch, Günter J, Ali Hortacsu, and Xiliang Lin**, “Prices and promotions in US retail markets: Evidence from big data,” *Quantitative Marketing and Economics*, 2021, *19* (3), 289–368.
- Hortacsu, Ali, Olivia R Natan, Hayden Parsley, Timothy Schweg, and Kevin R Williams**, “Organizational Structure and Pricing: Evidence from a Large U.S. Airline,” *The Quarterly Journal of Economics*, 2024, *139* (2), 1149–1199.
- Huang, Yufeng**, “Pricing Frictions and Platform Remedies: The Case of Airbnb,” June 2021.
- , **Paul B Ellickson, and Mitchell J Lovett**, “Learning to Set Prices,” *Journal of Marketing Research*, 2022, *59* (2), 411–434.
- Kraft, Andreas and Raghunath Singh Rao**, “Market effects of inattention: Theory and evidence from left-digit bias,” *JMR, Journal of Marketing Research*, February 2026, (00222437261427475), 00222437261427475.
- Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor**, “Heuristic Thinking and Limited Attention in the Car Market,” *American Economic Review*, 2012, *102* (5), 2206–2236.
- Lant, Theresa K.**, “Aspiration Level Adaptation: An Empirical Exploration,” *Management Science*, 1992, *38* (5), 623–644.
- Li, H and X Qiu**, “Heuristics in Self-Evaluation: Evidence from the Centralized College Admission System in China,” *Rev. Econ. Stat.*, 2023.
- List, John A, Ian Muir, Devin Pope, and Gregory Sun**, “Left-digit bias at Lyft,” *The Review of Economic Studies*, November 2023, *90* (6), 3186–3237.
- Maćkowiak, Bartosz and Mirko Wiederholt**, “Optimal Sticky Prices under Rational Inattention,” *American Economic Review*, 2009, *99* (3), 769–803.

- Matějka, Filip**, “Rationally inattentive seller: Sales and discrete pricing,” *The Review of Economic Studies*, 2015, *83* (3), 1125–1155.
- Misra, Kanishka, Eric M. Schwartz, and Jacob Abernethy**, “Dynamic Online Pricing with Incomplete Information Using Multiarmed Bandit Experiments,” *Marketing Science*, 2019, *38* (2), 226–252.
- Nevo, Aviv**, “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 2001, *69* (2), 307–342.
- Owen, Guillermo**, “Values of Games with a Priori Unions,” in “Lecture Notes in Economics and Mathematical Systems” Lecture Notes in Economics and Mathematical Systems, Berlin, Heidelberg: Springer Berlin Heidelberg, 1977, pp. 76–88.
- Repetto, Luca and Alex Solis**, “The Price of Inattention: Evidence from the Swedish Housing Market,” *Journal of the European Economic Association*, 2020, *18* (6), 3261–3304.
- Reyes, Germán**, “Coarse Wage-Setting and Behavioral Firms,” *IZA Discussion Paper 17039*, 2024.
- Romer, David**, “Do Firms Maximize? Evidence from Professional Football,” *Journal of Political Economy*, 2006, *114* (2), 340–365.
- Rothschild, Michael**, “A Two-Armed Bandit Theory of Market Pricing,” *Journal of Economic Theory*, 1974, *9* (2), 185–202.
- Russo, Daniel and Benjamin Van Roy**, “Learning to Optimize via Posterior Sampling,” *Mathematics of Operations Research*, 2014, *39* (4), 1221–1243.
- Salant, Yuval and Jörg L Spenkuch**, “Complexity and Satisficing: Theory with Evidence from Chess,” *Review of Economic Studies*, March 2026, *93* (2), 1296–1322.
- Sayag, Doron, Avichai Snir, and Daniel Levy**, “Price setting rules, rounding tax, and inattention penalty,” *arXiv [econ.GN]*, November 2024.
- Schulz, Eric, Nicholas T. Franklin, and Samuel J. Gershman**, “Finding Structure in Multi-Armed Bandits,” *Cognitive Psychology*, 2020, *119*, 101261.
- Schwartzstein, Joshua**, “Selective attention and learning,” *J. Eur. Econ. Assoc.*, 2014, *12* (6), 1423–1452.
- Shapiro, Bradley T, Günter J Hitsch, and Anna E Tuchman**, “TV advertising effectiveness and profitability: Generalizable results from 288 brands,” *Econometrica*, 2021, *89* (4), 1855–1879.
- Shapley, Lloyd S and Others**, “A value for n-person games,” 1953.
- Shorrocks, Anthony F.**, “Decomposition procedures for distributional analysis: a unified framework based on the Shapley value,” *The Journal of Economic Inequality*, 2013, *11* (1), 99–126.
- Simon, Herbert A.**, “A Behavioral Model of Rational Choice,” *The Quarterly Journal of Economics*, 1955, *69* (1), 99–118.
- Sims, Christopher A**, “Implications of Rational Inattention,” *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

- Snir, Avichai, Daniel Levy, and Haipeng (allan) Chen**, “End of 9-endings, price recall, and price perceptions,” *Economics Letters*, June 2017, *155*, 157–163.
- Sokolova, Tatiana, Satheesh Seenivasan, and Manoj Thomas**, “The Left-Digit Bias: When and Why Are Consumers Penny Wise and Pound Foolish?,” *J. Mark. Res.*, August 2020, *57* (4), 771–788.
- Strulov-Shlain, Avner**, “More than a penny’s worth: Left-digit bias and firm pricing,” *Review of Economic Studies*, September 2023, *90* (5), 2612–2645.
- Sutton, Richard S and Andrew G Barto**, *Reinforcement Learning, second edition: An Introduction*, MIT Press, November 2018.
- Thomas, Manoj and Vicki Morwitz**, “Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition,” *Journal of Consumer Research*, 2005, *32* (1), 54–64.

Online Appendix, Not For Publication

Learning and Limitations of Heuristic Pricing: Evidence from a
Reform

Avner Strulov-Shlain

A Data Appendix

Prices and their dynamics are the core focus of this paper. This appendix describes how prices are represented in each dataset and the cleaning procedures used to construct the final samples.

Table A1: Summary Statistics of Prices Data

Variable	ICC		CBS		StoreNext	
	Raw	Main	Raw	Main	Raw	Main
No. of Products	67	21	233	171	19	17
No. of Stores	659	143	155	99	295-580	295-573
No. of Chains	38	20	—	—	12-23	12-23
Sampling Periods	41	40	48	48	156	156
First Observation	02/2013	03/2013	01/2012	01/2012	01/2013	01/2013
Last Observation	02/2015	02/2015	12/2015	12/2015	12/2015	12/2015
Mean Weekly Units Sold	—	—	—	—	56.69	66.09
Mean Price	12.99	6.95	10.48	10.04	12.36	11.06
Price SD	1.59	0.89	1.88	1.3	6.2	4.56
Pre 99 Share	0.38	0.45	0.36	0.47	0.13	0.23
Post 90 Share	0.52	0.59	0.39	0.48	0.29	0.39
Post 00 Share	0.12	0.10	0.09	0.12	0.15	0.15
Observations	472,572	116,995	292,471	156,336	846,392	185,775

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. StoreNext is a market research firm collecting scanner data from retailers, including quantities and revenues, at the store-week-item level. These data are a merge of two datasets, and thus there is likely an overlap between stores and chains which have unique IDs in the two datasets.

A.1 ICC

ICC data are collected by surveyors who visit a set of predetermined stores with a product list; each product corresponds to a UPC. The surveyors record the shelf price of the item, and the total basket price was posted on the ICC website. The micro-data were generously provided to me by the ICC. The product list may change from week to week, but a core set of products remained constant. Surveyors visited stores every 2–4 weeks: every 4 weeks between January and September 2014 due to budget cuts, and every 2 weeks otherwise.

In January 2015 Israel passed the “food law,” mandating full price transparency starting May 2015 (Ater and Rigbi, 2019). This made the ICC’s collection redundant and led to its cancellation, which is why the ICC panel ends in February 2015.

An important feature is that prices are recorded for the first unit purchased. A promotion such as “second item 50% off” is not reflected in the recorded price, so the ICC prices mostly represent non-sale prices.

Price imputation and sample selection. The ICC data are not a balanced panel: the product list changes over time, stores enter and exit, and the sampling rate is inconsistent. I therefore clean the data toward balance in three steps. First, I restrict the sample to products that are sampled in all periods. Second, I restrict to stores sampled at least T times within each policy-relevant period, with $T = 4, 3$, and 5 for the pre-announcement (3/13–10/13), announcement-to-enactment (10/13–12/13), and post-enactment (1/14–2/15) periods respectively in the Main ICC sample; for the post-only analysis I require 15 observations in the post period and an observation in the final period (February 2015). Third, I balance the panel by completing missing price observations for a product in a store backwards. That is, if a product’s prices in three consecutive sampling periods are $\{3.9, \text{Missing}, 4.2\}$, I complete them to $\{3.9, 4.2, 4.2\}$. Completing backwards sets an upper limit on measured learning rates.³⁷

This cleaning sharpens the patterns and is necessary for exploiting the panel structure of the data, but does not qualitatively alter the main results. Figure A-5 in Appendix B reproduces Figure 2 using the raw ICC data, without sample selection or price imputation.

A.2 CBS

The second pricing dataset is collected by Israel’s Central Bureau of Statistics (CBS) to construct the Consumer Price Index. It is similar in nature to the ICC data: surveyors collect displayed prices for a list of products from stores every month. Three differences matter. Because the CBS aims to reflect a representative shopping basket, it collects an extensive share of its data from non-supermarkets (markets, specialty shops, and convenience stores); each product is sampled from a somewhat different set of stores; and a “product” is a product-type rather than a UPC. For example, “cottage cheese” may differ in size, manufacturer, and fat content across stores, and within a store over time. To make the data comparable to the ICC sample, I restrict it in the same way and keep supermarkets only.

Against these limitations, the CBS panel is balanced, sampled monthly from 2012 through the end of 2015, and covers an order of magnitude more products than the ICC data (171 product-types in the final sample). Summary statistics appear in Table A1. Most of the pricing-pattern analysis

³⁷Completing forward is a robustness test; results did not change in a meaningful way.

uses the ICC data, since its chain identifiers allow the between-chain analysis that is of central interest; the basic analysis is run on both datasets and gives very similar results.

A.3 StoreNext

The StoreNext data are scanner data recording store-product-period revenue and units sold, together with anonymous store and chain identifiers. Prices in these data are therefore quantity-weighted averages of the prices paid during the period, equal to revenue over units.

The sample combines two extracts. The first covers 10 UPCs from 295 supermarket stores over 2013–2015, of which six are usable: one UPC has a single price throughout the sample (canned corn), one is a product introduced only after the reform (400g hummus), and two are cottage cheeses with unusual pricing behavior (see Hendel et al., 2017). The second adds weekly data on 13 further products from 287 national supermarket stores over the same period, giving 19 products in total and 17 after cleaning. The two extracts likely cover overlapping stores, but store identifiers are unique to each source and cannot be matched across them, so the store and chain counts reported for StoreNext in Table A1 are upper bounds: the true number of stores lies between 295 and 580 and is likely much closer to 295, and the number of chains is likely closer to 12 than to 23. The merge can split one real chain across two identifiers but cannot pool distinct chains, so a coded chain is a subset of a true chain. StoreNext is the only research-accessible scanner data source for Israel, so I address the measurement problem through cleaning rather than through an alternative source.

This introduces measurement error. To see the issue, consider a true series of prices for a chocolate bar in one store on four consecutive days in 2015, where the price is revenue divided by units sold. On Sunday the price was 10.9025 (4 units sold), on Monday exactly 10.90 (1 unit), and on Tuesday 10.902308 (13 units). Sunday’s figure is an average of 3 units at 10.90 and 1 unit at 10.91; Tuesday’s of 10 units at 10.90 and 3 at 10.91. A same-day price difference of 1 Agorot is extremely unlikely, and 10.91 was not an admissible price in 2015. When the average is this close to 10.90 the true price is easy to guess, but prices also fluctuate by a few Agorot from day to day in ways that are unlikely to be genuine price updates and harder to resolve. For a sense of the magnitude: weekly data from a national US retailer show 90% of prices “to-the-cent” (Strulov-Shlain, 2023), while the *daily* Israeli data show only 55%.

I aggregate the data to the weekly level in order to merge it with a second excerpt covering 13 additional products, and keep records from 2013 through 2015.

Given the measurement issues, I round prices to the nearest admissible price—at the 1-Agora level in 2013 and the 10-Agorot level in 2014–2015. To be confident that the weekly price represents the price consumers actually paid, I keep observations whose price is admissible absent rounding (34.6% of prices) or that belong to a same-price spell of at least three weeks (35.8%), retaining

47.26% of the product-store observations that pass the balanced-sample filter described below.³⁸ A further 2 percentage points are lost in estimation to fixed-effect cells too small to identify and to observations where the leave-out chain price is undefined, so the demand regressions use about 45% of that panel. In the resulting main sample, 72% of observations have admissible prices and 84% belong to same-price spells of at least two weeks. Descriptives appear in the “StoreNext Main” column of Table A1. Two spell thresholds appear in the paper and serve different purposes: the three-week threshold above is an inclusion rule for the main scanner sample, while the “long price spells” sample used in some analyses of Section 3 applies a looser two-week restriction to exclude likely promotional activity.

Balanced-sample construction. To create a balanced sample of product-stores, I keep product-store pairs with any sales in at least 80% of total weeks, and in 90% of the weeks of the pre-reform year 2013. For each price sequence of a product in a store I then compute the length of same-price spells, treating two consecutive weeks as the same price if they differ by at most 2 Agorot—unless the change is between a .99 and a .00, which I treat as distinct prices.

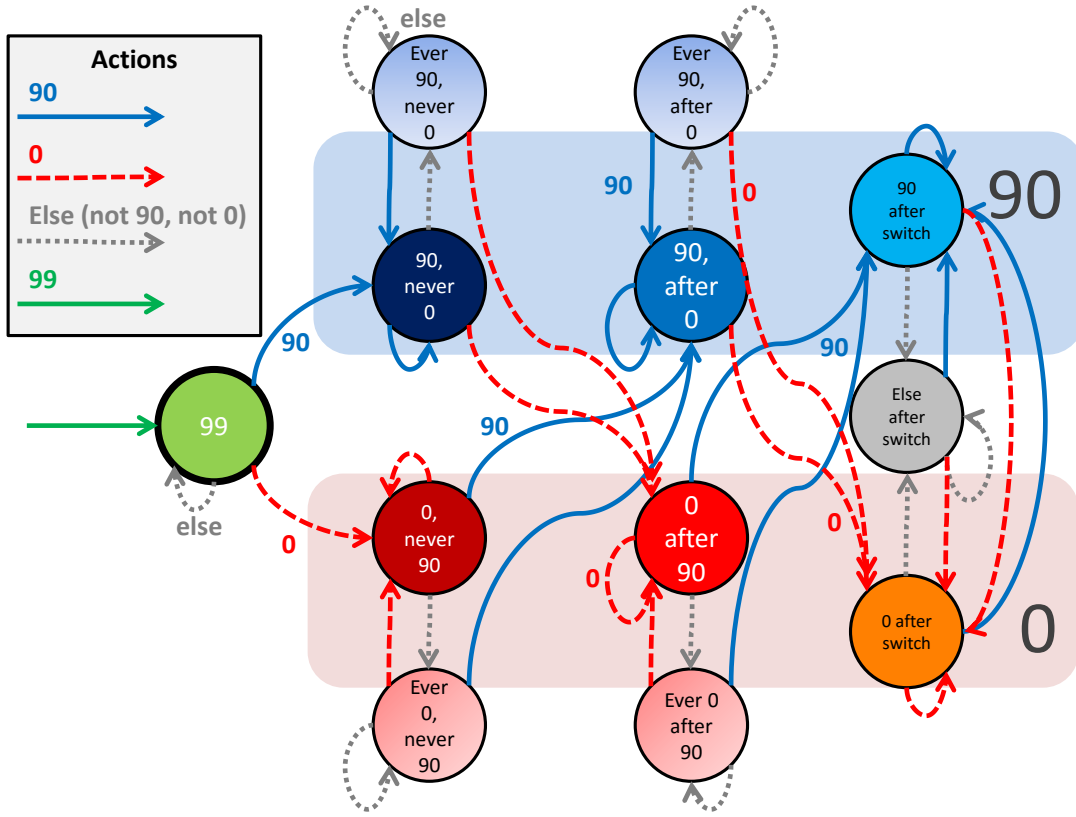
B Additional Figures and Pricing Patterns

B.1 Price-Ending Path Classification

Figure A-1 illustrates the state-machine used to classify price-ending paths at the product-store level, referenced in Section 3.

³⁸For same-price spells I code “the” price as the modal price of the spell.

Figure A-1: Automaton tracking price endings path

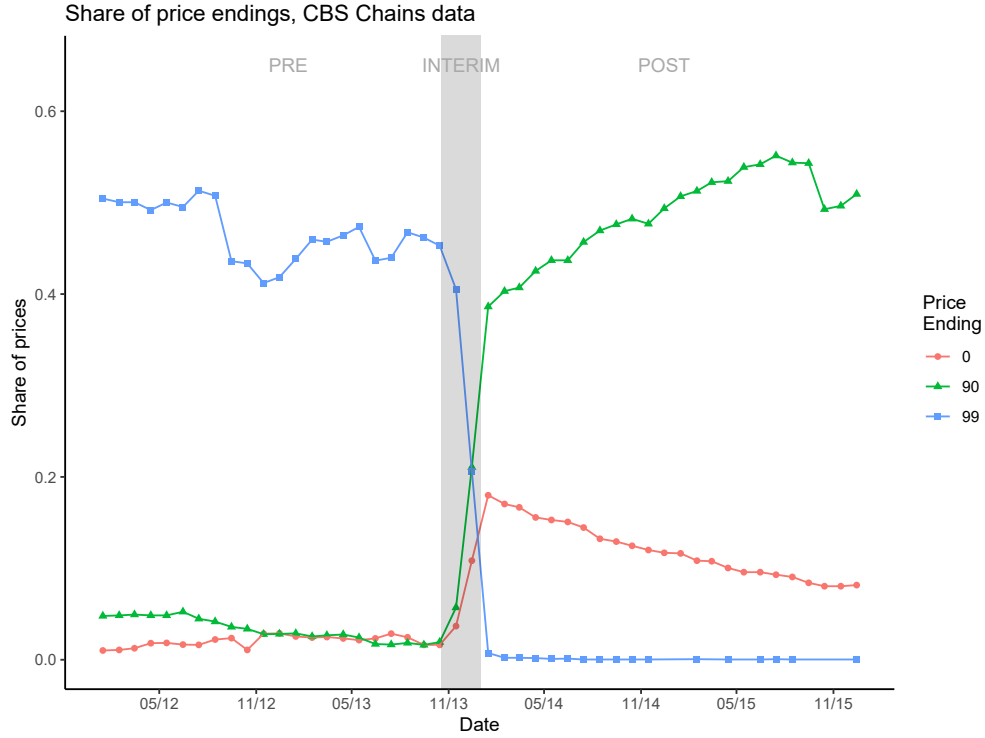


The figure illustrates a state-machine that assigns each observation into a “price-ending path” as a function of 4 possible price endings: 99, 90, 00, or other. For example, a price sequence of {4.99, 4.90, 5.00, 5.90} will have the corresponding states {99, 90 never 0, 0 after 90, 90 after switch}.

B.2 CBS replication of the main shares figure

Figure A-2 replicates Figure 2 in the CBS data. The patterns are qualitatively similar.

Figure A-2: Shares of price endings over time, CBS data

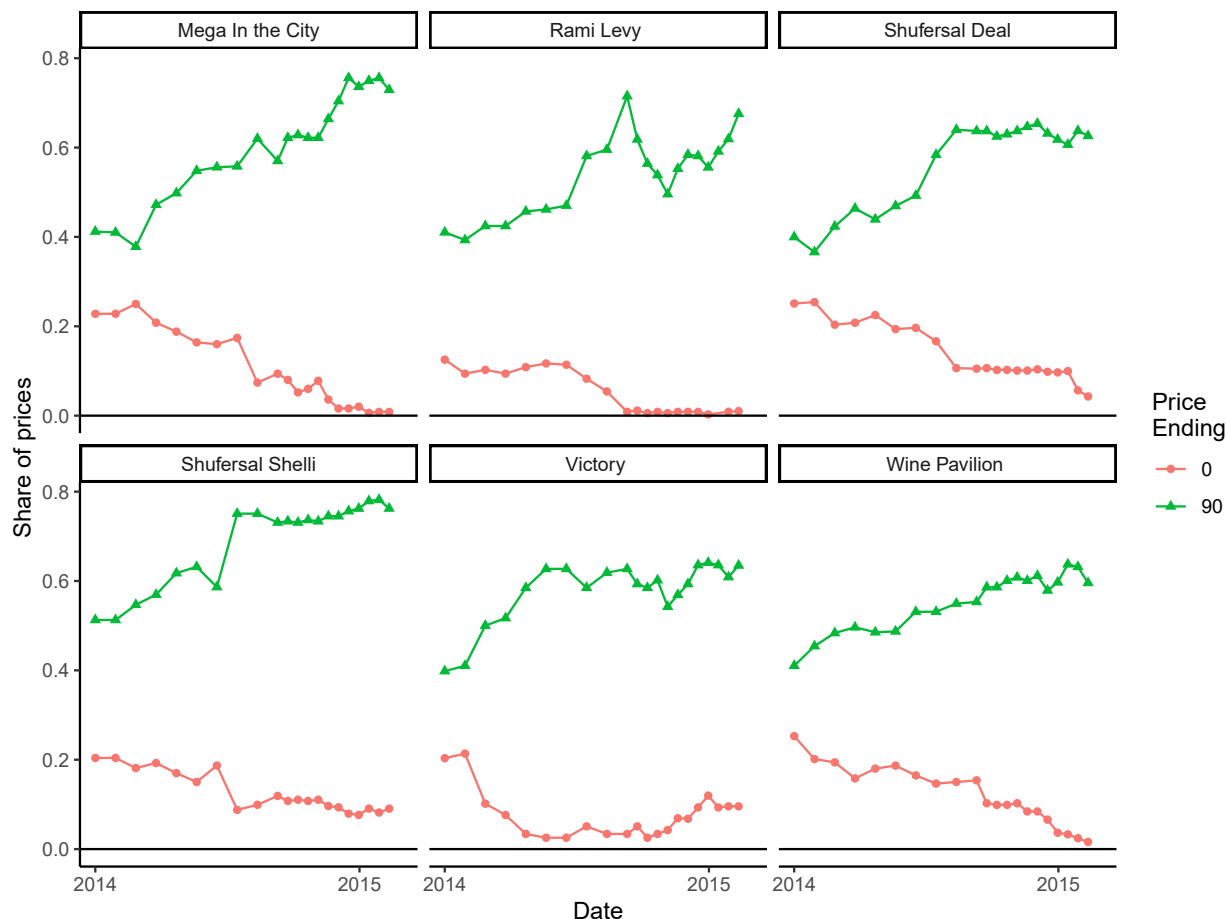


Same as Figure 2 but using the CBS data. Squares represent .99 prices, triangles .90 prices, and circles are .00 prices.

B.3 Price-ending shares by chain

Figure A-3 shows price-ending shares after the policy change for the six largest chains. For most chains, the .00 share drops in a stepwise fashion, consistent with chain-coordinated rollout rather than gradual store-by-store adjustment. The discrete drops support the interpretation in Section 3.3 that chains implemented product-specific corrections centrally; they do not by themselves locate where the underlying inference occurred.

Figure A-3: Shares of price endings post-reform by chain



The figure shows the shares of 90- and .00 prices across products and stores by sampling period, broken down by chain. Data show shares at the post-reform period (starting January 2014). Results are shown for the 6 biggest chains in the data.

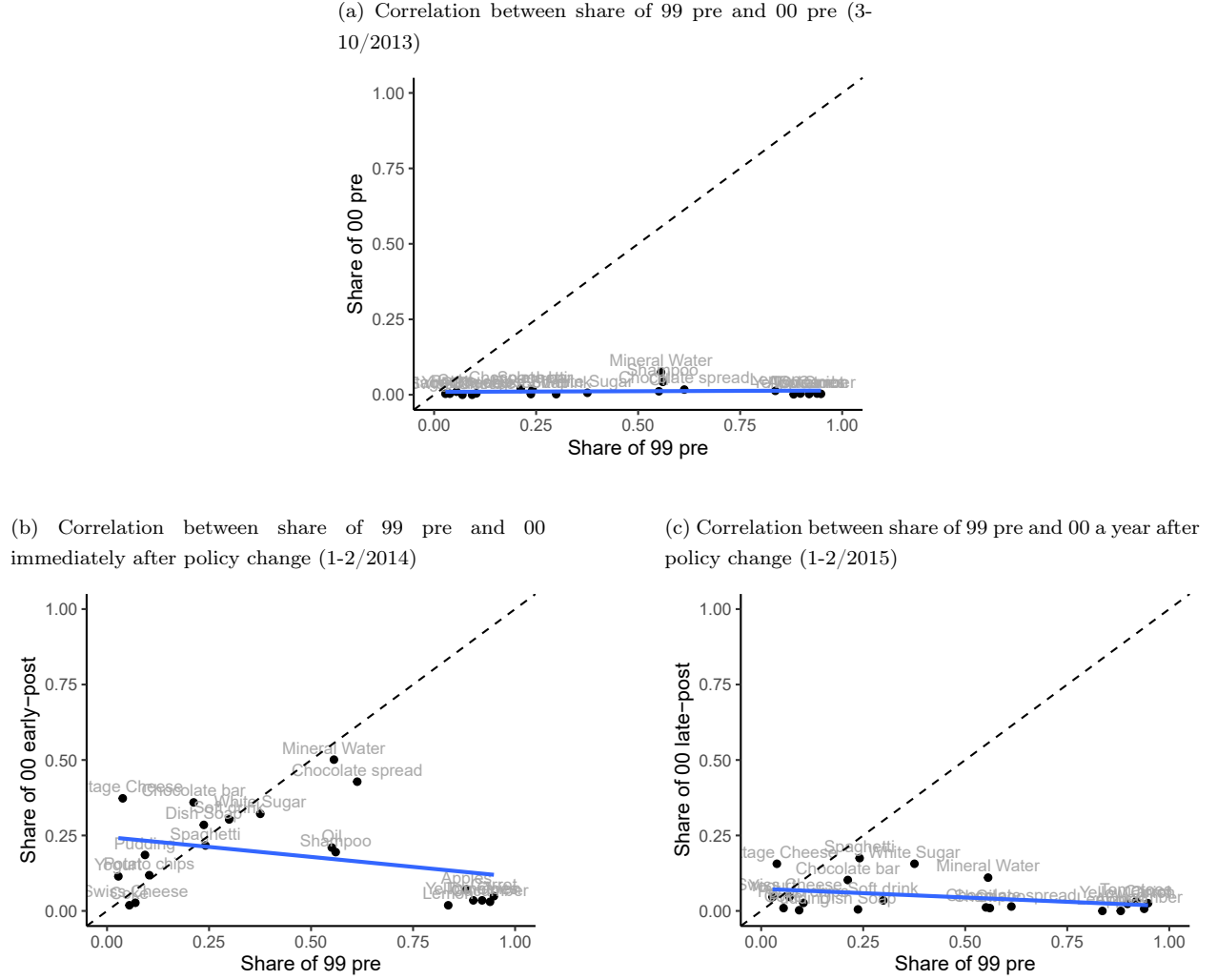
Triangles represent .90 prices and circles .00 prices.

B.4 Pre-reform 99-share vs. post-reform 00-share

Figure A-4 shows the relationship between the pre-reform share of .99s and the share of .00s before, immediately after, and a year after the reform, with each dot a product in the ICC data. The early-post panel exhibits the positive correlation that Section 3.4 reconciles with the binary first-transition regression's negative coefficient (see the predictor regressions in Appendix E). Panel (b) shows a *positive* correlation between a product's pre-reform .99-share and its early .00-share, which may appear to contradict the negative coefficient in the first-transition regressions of Appendix E. Yet, the two condition on different sets and are compatible. Products that relied most on .99 had the most prices forced to move at the reform, so more of their prices ended up at .00 *in levels*, which is what panel (b) plots. The regressions instead condition on the at-risk set: among the prices

that had to move, those same products were *less* likely to choose .00. The raw correlation reflects exposure to the reform; the regression coefficient reflects the conditional propensity.

Figure A-4: Correlation between pre-shares of 99 and shares of 00 across periods.

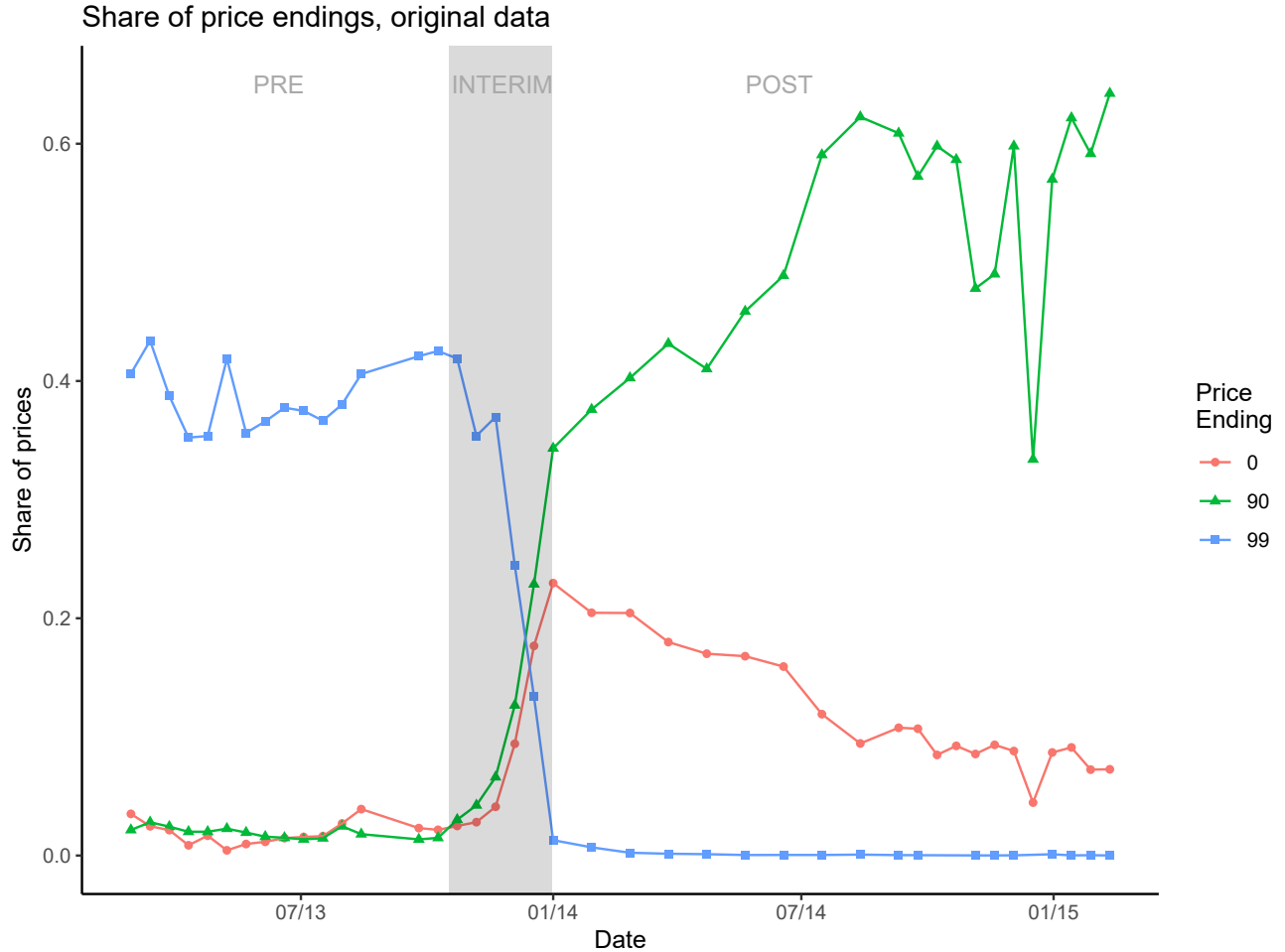


The figures show shares of .00 prices at different periods against .99 prices at the pre-period, by product. Each dot is a product in the ICC data.

B.5 Raw-data robustness

Figure A-5 shows the main pricing patterns using the raw ICC sample, without the panel-balancing cleaning described in the Data Appendix.

Figure A-5: Price ending patterns using the full ICC sample



The figure shows shares of 99-, 90-, and .00 prices over time using the raw ICC data without sample selection or price imputation. The patterns are qualitatively similar to those in Figure 2.

C Demand specification — the 10-column table

Table 1 reports the demand specification in Equation 3 across three instrumental-variable strategies. The first column block (columns 1–6) varies sample restrictions under the chain-interacted Hausman price IV. The second block (columns 7–8) replaces the Hausman IV with leave-out chain shares of each ending as instruments for κ^{90} and κ^{00} , with a pooled log P instead of chain-interacted prices. The third block (columns 9–10) combines both IVs.

From gaps to $\hat{\theta}$. The implied bias parameter inverts the perceived-price ratio at the .00 boundary. With bias θ , the perceived ratio of an integer price x (a .00 ending) to the 10-Agorot-lower .90 price is $x/(x-0.1-0.9\theta)$, so the log-demand gap at elasticity ϵ is $|\epsilon| \ln[x/(x-0.1-0.9\theta)] \approx |\epsilon| (0.1+0.9\theta)/x$.

Inverting, $\hat{\theta} = [x(1 - e^{-\text{gap}/|\epsilon|}) - 0.1] / 0.9$. Evaluated at the mean .00 price in 2014 ($x = 10.3$) and the estimated elasticity ($\epsilon = -2.01$), the six Hausman-IV gaps (0.065–0.112) imply $\hat{\theta}$ of 0.28, 0.34, 0.34, 0.51, 0.25, and 0.38 for columns 1–6: five of the six lie in $[0.25, 0.38]$, and the most restrictive filter (column 4) is the outlier at 0.51. The ending-IV gaps (0.15–0.23) would imply $\hat{\theta}$ of 0.9 or more at the same inputs; the text therefore uses the conservative Hausman set, and the ending-IV columns are better read as bounds on the demand response than as point estimates of θ .

Robustness to Jewish-holiday timing. A natural concern is that Hebrew-calendar-driven seasonal variation might contaminate the estimated penalty. Adding seven year-specific Jewish-holiday window indicators to columns 4 and 7 changes the coefficients by at most 0.003 in absolute value, and the 90-vs-00 gap by less than 0.001. The heavy fixed-effect saturation absorbs essentially all of the variation that holiday flags could capture.

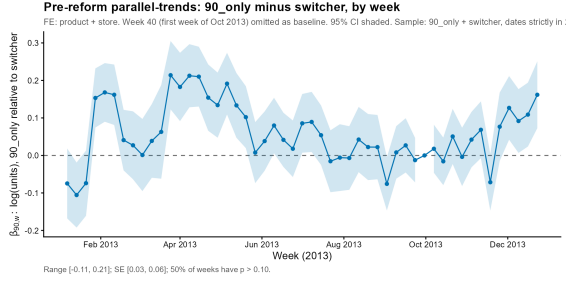
D Pre-reform parallel-trends tests

One concern with the demand specification in Section 3 is that product-stores that eventually used .00 endings vs. .90 endings may have had differential demand trends in 2013. To test this, I run two coefficient-plot specifications on 2013-only data with a class-by-week interaction and product×store fixed effects, store-clustered standard errors. Under the “eventual ending” classification (.90_only vs. switcher), the weekly coefficients range over $[-0.11, +0.21]$ and fluctuate around zero with no drift (Figure A-6a); because both cells are large, however, half of the weekly coefficients differ from the baseline week at the 10% level: scattered, precisely estimated level differences rather than a visible trend. Under the “first-transition” classification (first_00 vs. first_90, the classification used in the paths figure), 88% of weeks have $p > 0.10$ and the coefficient range is $[-0.36, +0.22]$, with wider confidence intervals from the smaller first_00 cell (Figure A-6b).

Formal tests agree with this reading. A linear differential trend (class × week index) is small and insignificant under both classifications: -0.0008 log points per week (SE 0.0007, $p = 0.19$; about -0.04 over the year) for eventual ending, and $+0.0015$ (0.0028, $p = 0.58$) for first transition. A joint Wald test that all class-by-week interactions are zero is rejected in both ($F = 4.9$ and 8.9 , $p < 0.001$): individual week-to-week level differences are precisely estimated, but they do not accumulate into a trend. The classes also differ in pre-period levels (eventual ending $+0.11$, $p < 0.001$; first transition -0.36 , $p = 0.009$), which the product×store fixed effects in every demand specification absorb. Neither classification shows a differential trend. The sharper answer to the underlying selection concern is in any case the instrumented estimates: the 90-vs-00 gap *widens* under the ending IVs (Table 1), the opposite of what adverse selection into .00 would produce.

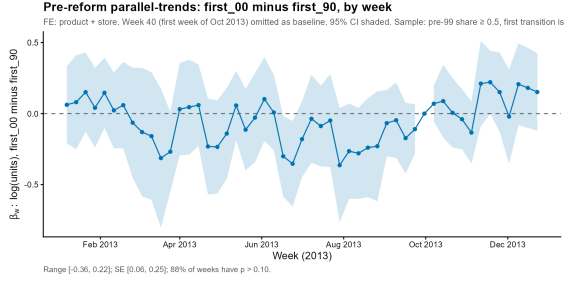
Figure A-6: Pre-reform parallel-trends tests, two classifications

(a) By eventual post-reform ending (.90_only minus switcher)



(b) By first non-.99 transition (first_00 minus first_90)

Both



panels regress $\log(\text{units}) \sim \text{class} \times \text{week-of-2013} \mid \text{product} + \text{store}$, store-clustered SE, week 40 (first week of October 2013) as baseline. 95% CI shaded. Left panel: 889 + 732 product-stores. Right panel: 40 + 97 product-stores (smaller first_00 cell $\rightarrow$ wider CI).

E Predictors of first transition to .00

The variance decomposition in Section 3.4 attributes 55% of the explained variance in the binary first-transition outcome to product identity in-sample (75% under cross-validation). This appendix reports the underlying regressions and decomposition.

E.1 Stacked cross-dataset regression

I run the first-transition specification on three datasets (ICC, CBS, StoreNext) and stacked across them with dataset fixed effects. The outcome is the binary indicator $\text{first_00} \in \{0, 1\}$ defined for product-stores with pre-reform .99-share ≥ 0.5 whose first non-.99 ending is in $\{.00, .90\}$. The predictor is the continuous share of .99 endings in the pre-reform period. The coefficient on **share_99_pre**: -0.265 (stacked, no fixed effects, $p < 0.001$); -0.325 (stacked, store fixed effects, $p < 0.001$); -0.531 (stacked, chain fixed effects, $p < 0.001$); but $+0.023$ (stacked, product fixed effects, not significant). The negative coefficient vanishes once product fixed effects are added. The same reversal holds dataset by dataset, not only in the pooled fit. In the ICC sample alone the coefficient on **share_99_pre** is -0.43 without product fixed effects and $+0.05$ (not significant) with them. The negative association is therefore a between-product regularity present in each sample rather than an artifact of pooling across datasets with different product definitions.

E.2 Full predictor regression

For the StoreNext sample augmented with pre-reform sales, the predictor regression yields the following coefficients (store-clustered standard errors). Without product fixed effects: pre-

reform mean price +0.051 ($p < 0.01$); **share_99_pre** -1.19 ($p < 0.05$); $\log(n_{\text{stores chain}}) -0.11$ (not significant); pre-reform $\log(\text{sales})$ 0.020 (not significant). With product fixed effects: pre-reform mean price not identifiable (saturated); **share_99_pre** -2.81 ($p < 0.01$); other coefficients absorbed or not significant. With product and chain fixed effects, none of the predictors is significant. For the ICC sample with 5-km competitor count, none of the predictors survives product fixed effects either.

E.3 Shapley-Owen variance decomposition with cross-validation

A Shapley-Owen decomposition (Shapley and Others, 1953; Owen, 1977) uniquely allocates R^2 across the predictor blocks by averaging marginal R^2 contributions across all $n!$ orderings; the same construction is standard in inequality decomposition (Shorrocks, 2013) and in the variance-decomposition literature. In the ICC sample of 2,887 product-stores at risk, the predictors account for an in-sample R^2 of 0.468 and a 10-fold CV R^2 of 0.333; on the CBS sample of 3,540 product-stores, 0.530 and 0.389. Two value functions are reported: in-sample R^2 and held-out R^2 under 10-fold cross-validation. The CV value function fits each predictor subset on 9 folds and predicts the held-out fold; observations whose FE level was not seen during training are assigned the training-sample mean of the outcome. By construction, each set of Shapley values sums (to machine precision) to the corresponding full-model R^2 .

The candidate predictors are: product fixed effects, store fixed effects, chain fixed effects (ICC only; CBS has no chain identifier), the continuous pre-reform mean price, and the continuous pre-reform .99 share. Outcomes are run separately: the binary first-transition indicator (at-risk panel), the continuous 2013 price (full panel), and the continuous 2014 price (full panel).

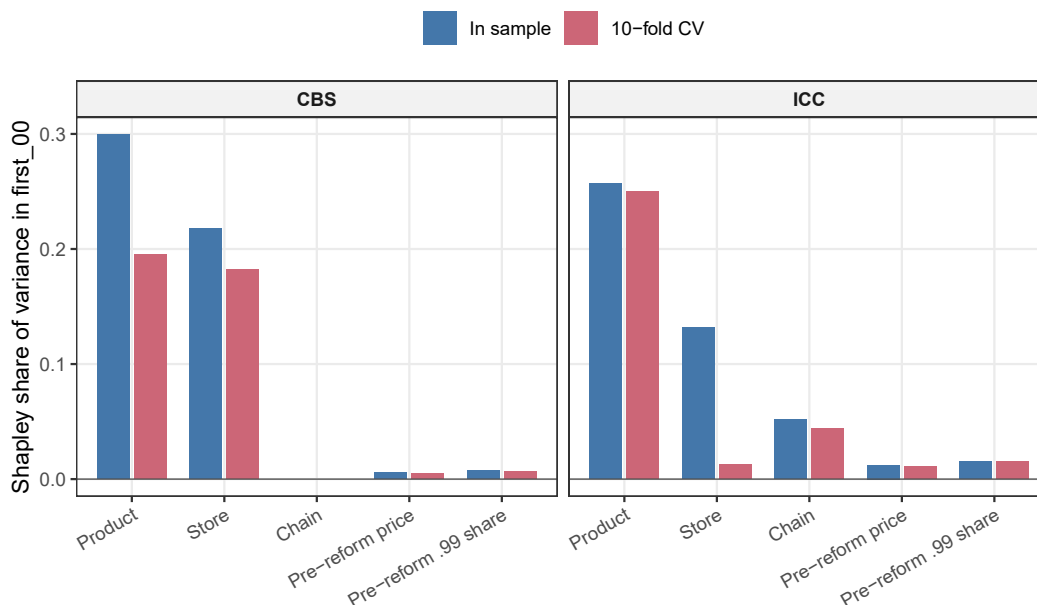
Table A2: Shapley-Owen variance decomposition: in-sample and 10-fold CV

Dataset	Outcome	N	R^2_{in}	R^2_{CV}	Product		Store		Chain		price_pre		share_99_pre	
					in	CV	in	CV	in	CV	in	CV	in	CV
ICC	first_00	2,887	0.468	0.333	0.257	0.250	0.132	0.013	0.052	0.044	0.012	0.011	0.015	0.015
ICC	price_2013	8,741	0.991	0.990	0.986	1.017	0.003	-0.027	0.001	-0.001	—	—	0.001	0.000
ICC	price_2014	8,741	0.990	0.989	0.498	0.508	0.002	-0.018	0.001	-0.001	0.489	0.499	0.000	0.000
CBS	first_00	3,540	0.530	0.389	0.300	0.195	0.218	0.182	n/a	n/a	0.006	0.005	0.007	0.007
CBS	price_2013	8,039	0.727	0.694	0.668	0.659	0.058	0.034	n/a	n/a	—	—	0.001	0.001
CBS	price_2014	8,039	0.797	0.760	0.372	0.356	0.045	0.023	n/a	n/a	0.379	0.380	0.001	0.000

Each pair (in, CV) is the in-sample and 10-fold cross-validated Shapley share. Shapley shares within a row sum to the row's

full-model R^2 up to numerical precision. Negative CV shares (e.g., Store under **price_2013** for ICC) indicate predictors that are absorbed in-sample but actively hurt out-of-sample prediction because the variation they fit is already captured by another predictor more efficiently. “**price_pre**” is omitted as a predictor when the outcome is **price_2013** (perfect collinearity). CBS has no chain identifier.

Figure A-7: Shapley-Owen variance decomposition of the first-transition outcome



Shapley shares of variance in the binary first-transition indicator (.00 vs. .90) at the product-store level across five candidate predictors: product FE, store FE, chain FE (ICC only), pre-reform mean price, and pre-reform .99 share. Blue bars are the in-sample Shapley values; red bars are the 10-fold cross-validated Shapley values. The full table including the price-2013 and price-2014 outcomes is in Table A2.

Sequential (first-claim) decomposition. The Shapley shares average each block’s marginal contribution over all orderings, which divides variance that more than one block could explain. A complementary way to bound the observed covariates’ explanatory power is to give them first claim: enter `price_pre` and `share_99_pre` alone, then add the fixed effects. Ordered first—their most favorable position—the two covariates reach an in-sample R^2 of 0.06 in ICC and 0.03 in a stacked fit across datasets, against a further 0.22 from adding product fixed effects.³⁹ Because a block’s first-position R^2 is an upper bound on its Shapley share, this confirms rather than qualifies the Shapley result: the observed covariates are weak predictors of the first-transition outcome even under the ordering most generous to them, and their Shapley shares in Table A2 (0.012 and 0.015 in ICC) are correspondingly smaller.

Substantive takeaways. The cross-validated decomposition sharpens the conclusion in Section 3.4: product identity is the dominant predictor across all six (dataset $\times$ outcome) cells, and is robust to cross-validation (in-sample and CV Shapley shares agree to within a few percentage points). Store identity in ICC is a substantial in-sample contributor to `first_00` (0.132) but collapses to 0.013 out-of-sample: the in-sample Store “signal” is essentially overfit, driven by thin within-store data (~ 12 observations per store on the at-risk sample). CBS has fewer stores with more observations each,

³⁹The stacked fit is not shown in Table A2, which reports ICC and CBS separately.

and Store survives CV (0.218 vs. 0.182). The continuous covariates `share_99_pre` and `price_pre` are negligible predictors of `first_00` once product is in the model. For `price_2014`, lagged price (`price_pre`) accounts for about half the explained variance in both datasets (~ 0.49 in ICC, ~ 0.38 in CBS); within-product-store price persistence is strong even across the reform discontinuity. The reported Shapley shares sum exactly to the full-model R^2 in every cell, satisfying the efficiency property.

E.4 Product-level heatmaps of the transition

To make the product-level organization of the transition concrete, a store-by-week heatmap of the price-ending state for a single product shows, at a glance, the chain-synchronized execution (stores of a chain flipping together), the product-specific choice of ending, and the one-way .00-to-.90 correction. Figure A-8 shows four contrasting products. Mineral water is the heaviest .00 adopter (119 adopting product-stores) and corrects to .90 over 2014–15; white sugar also adopts .00 heavily (90 product-stores) but mostly keeps it (32 correct), a never-recoverer case consistent with the selection evidence in Appendix J.3; cucumber has the highest pre-reform .99-reliance (0.94) and goes essentially straight to .90; and Coca-Cola, with almost no pre-reform .99-reliance (0.05), barely moves its endings at all. Heatmaps for all 21 ICC products—price levels, price endings, and within-product demeaned prices—are available in the replication package; the contrast between products that moved to .00 and products that went directly to .90 is visible product by product rather than chain by chain.

Figure A-8: Price-ending states by store and week, selected products

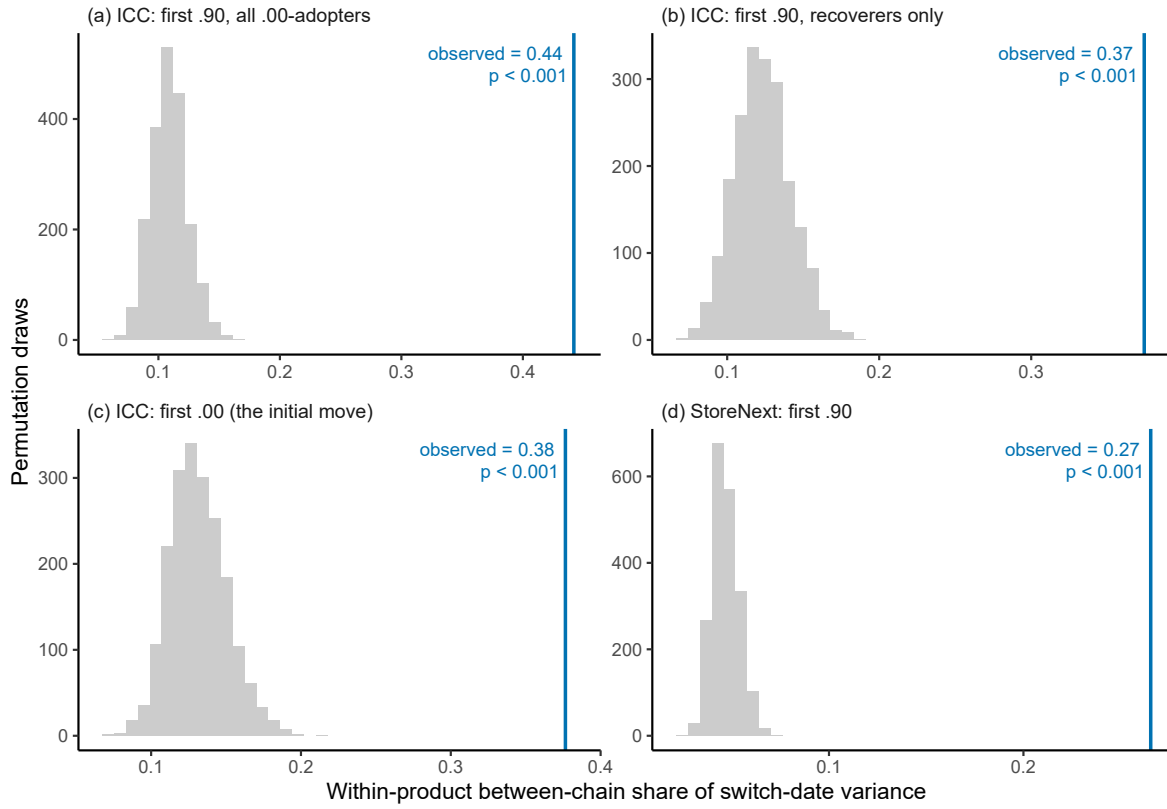


Each panel shows one product; rows are stores (grouped by chain, chain bands labeled at left), columns are biweeks, and colors are the price-ending state (.99, .90, .00, other). Dashed vertical lines mark the reform announcement (October 2013) and enactment (January 2014).

F Chain synchronization of switch timing

This appendix reports the permutation test behind the synchronization result in Section 3.5. For each timing outcome, the statistic is the within-product between-chain share of the variance in the first-occurrence date. The null distribution reshuffles chain labels among a product’s stores (2,000 draws), holding the dates fixed: it answers how much between-chain share would arise if stores reached the ending on chain-independent schedules. Figure A-9 shows four versions. Panel (a) is the version in Section 3.5: the first-.90 date among all 897 .00-adopters, with never-recoverers assigned the end of the sample—an observed share of 0.44 against a null centered at 0.11. Panel (b) restricts to the 704 product-stores that recover (0.37). Panel (c) applies the same test to the first-.00 date: the initial move was also executed on chain schedules (0.38). Panel (d) replicates the recovery result in the StoreNext scanner data (0.27 against a null centered at 0.05). All four reject chance at $p < 0.001$. Panel (c) also bears on how much information the timing of the initial move carries. The moves into .00 were concentrated at the reform’s onset and near-simultaneous across products, executed as part of the general re-pricing: within a product, a chain’s stores mostly moved together. The few later moves into .00 are scattered product-store events. Timing therefore distinguishes chains from one another but says little about *which* products received .00—that structure is between products (Appendix E).

Figure A-9: Chain synchronization of switch timing: permutation tests



Each panel shows the permutation-null distribution (2,000 draws; chain labels reshuffled among a product's stores) of the within-product between-chain share of first-occurrence-date variance, against the observed share (vertical line). Panels (a)–(c): ICC, top seven chains; panel (d): StoreNext. Panel (a) keeps every .00-adopter and assigns never-recoverers the end of the sample; panels (b) and (d) restrict to product-stores that reach .90 after .00.

G Cross-product spillover decomposition

As a complementary, product-level view of the chain-coordinated recovery documented in Section 3.5, I decompose switching into four spillover channels. The dependent variable is the binary indicator that a (chain $\times$ store $\times$ product) cell that has hit .00 makes its first transition to .90 in the focal biweek. The four predictors count, for each focal observation, the number of *other* products experiencing a switch event in the indicated cell.

Three of the four channels are significant at $p < 0.001$, univariately and jointly. The largest is same-biweek within-store at +6.3 percentage points (combined spec), followed by within-store recent at +3.9 and same-biweek chain at +0.6; the within-chain *recent* channel is small and not significant (+0.15, $p = 0.17$) once the lag windows are indexed to calendar biweeks rather

than observed sampling periods.⁴⁰ The same-biweek within-store channel reflects store managers re-pricing several products at once (the product-by-store execution noted in Section 3.3), while the same-biweek chain channel reflects synchronized rollouts. Chain influence operates through same-week coordinated moves rather than gradual within-chain diffusion, consistent with chain-coordinated, feedback-driven correction rather than reversion that is independent across a product’s stores.

H Detection: benchmark precision and reversion

This appendix reports the estimates behind the detection result in Section 3.5, that the demand shortfall at .00 predicts faster reversion most where the firm’s pre-reform demand benchmark is precise, together with an ex-ante placebo for the precision measure. The benchmark sample is the 514 .00 episodes in the cleaned StoreNext panel with at least two shortfall observations in the first four weeks and at least six pre-reform observations; 88% eventually switch to .90. Column 0 of Table A3 drops the pre-reform observation requirement and so runs on all 536 episodes. Because the shortfall is constructed from observations during the first four weeks of the episode and censored at the switch, the hazard estimates are predictive associations rather than a design that randomly assigns an ex-ante signal. *Benchmark precision* is the negative log of the standard deviation of a product-store’s pre-reform demand residuals (from the same baseline model used to construct the shortfall), standardized. *Just-below adherence* is the pre-reform share of endings with a 9 tens-digit, standardized; *data richness* is the log count of pre-reform observations, standardized.

Table A3 shows the pattern in four steps. Column 0 is the feedback effect on the full episode sample, before any precision requirement: the shortfall alone carries a hazard-ratio of 1.11 per standard deviation ($p = 0.022$, $N = 536$). Columns 1–2 report the Cox proportional-hazards estimates on the benchmark sample: the shortfall and precision enter additively in column 1, and their interaction in column 2 is positive and significant (0.190, $p = 0.001$; a hazard ratio of 1.21 per standard deviation of each). Column 3 runs the horse race: interacting the shortfall simultaneously with precision, just-below adherence, and data richness, only the precision interaction survives ($p = 0.002$); the adherence and richness interactions are essentially zero. Column 4 replicates the interaction in a linear-probability model of switching within 16 weeks. First, precision’s *main* effect is negative: firms with a precise benchmark are, absent a bad signal, slower to abandon .00. Precision therefore matters through the signal rather than as a general propensity to switch. Second, the moderator is the precision of the benchmark rather than the share of just-below pricing itself; what matters is how well the firm can tell a shortfall from noise.

Columns 5–6 test the remaining confound: that benchmark precision proxies unobserved firm quality—attentive, well-run operations that both avoid mistakes and correct them faster. Under

⁴⁰ICC sampling ran at roughly four-week intervals during January–August 2014, so windows defined over observed periods would span twice the calendar time there.

that account, precise product-stores should have avoided .00 at the reform to begin with. The first-transition sample shows no such protection. Among product-stores predominantly at .99 pre-reform whose first post-reform ending is .00 or .90 ($N = 281$), the precision coefficient on choosing .00 is zero without controls (column 5) and *positive* within product and chain (column 6), the opposite sign to the quality account. The positive coefficient leans on a single product (leave-one-product-out estimates span -0.004 to $+0.062$), and precision does not predict when a product-store first moved off .99 ($+0.1$ days per standard deviation, $p = 0.81$). A precise benchmark mattered only after the fact, once realized demand at .00 could be compared against it. In the language of Section 4.1, the evaluation noise behind the initial mistake is a property of the untried option rather than of the firm's knowledge of its familiar price.

Table A3: Demand shortfall $\times$ benchmark precision: the .00 $\rightarrow$.90 reversion and the ex-ante placebo

	Cox PH (hazard of first .90)				LPM (switch $\leq$ 16w)	LPM (first .00)	
	(0)	(1)	(2)	(3)	(4)	(5)	(6)
Shortfall (z)	0.105 (0.046)	0.098 (0.047)	0.120 (0.053)	0.091 (0.057)	0.045 (0.012)		
[p-value]	[0.022]	[0.035]	[0.023]	[0.108]	[0.002]		
Benchmark precision (z)		-0.233 (0.067)	-0.224 (0.067)	-0.155 (0.074)	-0.004 (0.023)	0.002 (0.057)	0.049 (0.014)
[p-value]		[0.001]	[0.001]	[0.035]	[0.868]	[0.977]	[0.004]
Shortfall $\times$ precision			0.190 (0.057)	0.178 (0.056)	0.061 (0.014)		
[p-value]			[0.001]	[0.002]	[0.001]		
Just-below adherence (z)				0.220 (0.074)			
[p-value]				[0.003]			
Data richness (z)				0.224 (0.072)			
[p-value]				[0.002]			
Shortfall $\times$ adherence				0.002 (0.055)			
[p-value]				[0.975]			
Shortfall $\times$ richness				0.032 (0.068)			
[p-value]				[0.644]			
Product + chain FE	—	—	—	—	Yes	No	Yes
<i>N</i>	536	514	514	514	514	281	281

Columns 0–3: Cox proportional-hazards coefficients on the weeks from a product-store’s first .00 to its first subsequent .90

(right-censored at the last observation), with product strata and store-clustered robust standard errors in parentheses; the hazard ratio is $\exp(\text{coef})$. Column 0 uses the full episode sample (at least two shortfall observations in the first four weeks);

columns 1–4 add the benchmark requirement (at least six pre-reform observations), which drops 22 episodes. Column 4: linear-probability model of switching within 16 weeks with product and chain fixed effects and chain-clustered standard errors.

All regressors are standardized. Exact p-values in brackets. The shortfall is the average residual demand gap relative to .90-priced chain peers over the first four weeks of the .00 episode (Section 3.5); benchmark precision is $-\log(\text{SD})$ of the product-store’s pre-reform demand residuals. Because a noisier demand history also makes the shortfall itself noisier, the behavioral and statistical readings of the interaction are not separable; both point to benchmark precision as what paces the correction. Columns 5–6 (ex-ante placebo): linear-probability models of a product-store’s *first* post-reform ending being .00 rather than .90, on the at-risk sample of product-stores with a pre-reform .99 share of at least one half and a measurable benchmark ($N = 281$; 16 products, 11 chains). Column 5 has no fixed effects; column 6 adds product and chain fixed effects; standard errors are chain-clustered (store-clustered errors and within-product randomization inference are similar: column 6 SE 0.015, randomization $p = 0.009$). The positive column-6 coefficient has the opposite sign from what a firm-quality account predicts, and it rests on a single product. Leave-one-product-out estimates span -0.004 to $+0.062$. In the fixed-effects row, the Cox columns are marked “—” because they stratify by product instead.

I Recovery ordering across products

This appendix reports the evidence behind the ordering result in Section 3.5: which products re-establish .90 first, and whether that order matches what a model-based firm would do.

The recovery has a systematic product-level order. Products that relied most heavily on just-below pricing before the reform recover faster and more fully: across products, the rank correlation between the pre-reform .99-share and the mean first-.90 date after a .00 is -0.55 ($p = 0.011$) when every .00-adopter is included, and -0.80 ($p < 0.001$) among the subset of product-stores that eventually change to .90. In a suggestive regression of first-.90 timing on standardized product characteristics ($N = 21$, adjusted $R^2 = 0.77$), the .99-share coefficient is -8.8 weeks per standard deviation ($p < 0.001$), about twice the price coefficient (-4.8 , $p = 0.002$), with pre-reform price-change frequency insignificant ($+0.3$, $p = 0.83$).

The ordering an optimizer would follow is not supported, though the test is asking a lot from the data. If firms corrected where the mistake was costliest, products for which .00 is dominated over the widest range (the lowest threshold θ_0) should recover first. However, the rank correlation between the recovery order and θ_0 is $+0.19$ ($p = 0.40$) when θ_0 is computed at a common elasticity in the ICC data, and $+0.07$ ($p = 0.81$) from each product's *estimated* demand elasticity in the scanner data. With 21 products these correlations are imprecise, so the sharper statement is that .99-reliance dominates the threshold ordering, not that the latter is rejected.⁴¹ The immediate-adoption version of the test points the same way: the product-level share of at-risk product-stores whose *first* post-reform ending was .00 is unrelated to θ_0 (rank correlations $+0.01$, $p = 0.96$, ICC; $+0.07$, $p = 0.80$, scanner data at product elasticities) but falls with pre-reform .99-reliance (-0.59 , $p = 0.022$) echoing the propensity result of Section 3.4.

J Demand stability — additional tests

J.1 Period-interacted gap under the ending-IV specification

Splitting the post-reform period into 2014H1, 2014H2, and 2015 and re-estimating gap-by-period under the endings-IV specification gives $+0.198$ (SE 0.057, $p < 0.001$) in 2014H1, $+0.161$ (0.067, $p = 0.016$) in 2014H2, and -0.025 (0.119, $p = 0.83$) in 2015. Pairwise, the 2014H1 and 2014H2 gaps do not differ ($+0.037$, $p = 0.64$), and the noisy 2015 gap differs significantly from neither ($p = 0.19$ vs. 2014H2; $p = 0.10$ vs. 2014H1). The 2015 cell is imprecise because, by then, few product-stores still use .00, and those that remain are a selected set (see Appendix J.3).

⁴¹Adding θ_0 to the timing regression does not change this. The .99-share coefficient is unchanged (-8.8 , $p < 0.001$); the common- ϵ θ_0 enters with a conditional coefficient of -3.2 ($p = 0.04$) whose sign is the *opposite* of the threshold prediction. Since θ_0 at a common elasticity is a deterministic decreasing transform of the pre-reform price, that term picks up price curvature rather than threshold content. The product-elasticity version, which varies independently, contributes nothing ($+2.0$, $p = 0.55$).

J.2 Elasticity stability: a structural-break test

The demand *elasticity* is likewise stable around the reform. On the cleaned StoreNext panel ($N = 218,129$), I estimate $\log Q$ on the ending dummies and $\log P$ interacted with four periods (2013, 2014H1, 2014H2, 2015), with product-store and month-of-year $\times$ product fixed effects, instrumenting $\log P$ with the period-interacted leave-out chain average price and clustering by store. The period elasticities are -2.01 (SE 0.04), -2.00 (0.04), -2.01 (0.04), and -2.02 (0.05); the OLS variants are -1.89 to -1.91 . Because the period estimates are extremely precise and highly correlated, a joint Wald test rejects exact equality ($\chi^2(3) = 37.2$, $p < 0.001$), with pairwise differences of 0.01–0.03. The rejection has no economic content: the largest difference is about 1% of the elasticity’s level, and no calculation in the paper (the θ_0 thresholds, the implied $\hat{\theta}$, or the forgone-profit series) changes when the elasticity moves from -2.00 to -2.03 .

J.3 2013-trained out-of-sample residual test

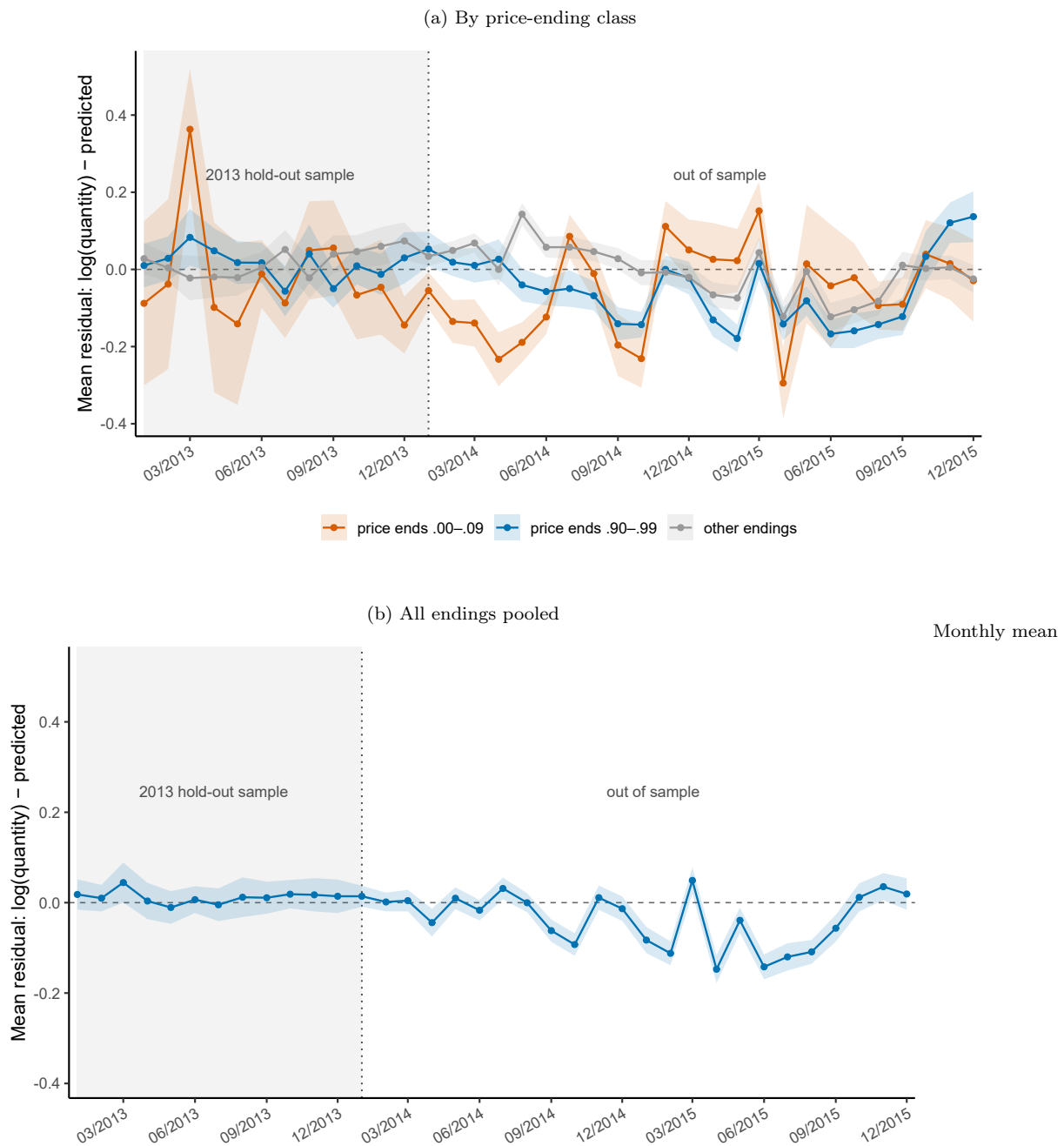
A different way to ask the demand-stability question is to train a demand model on 2013 and project it forward. I randomly split 2013 into 80% training (54,987 observations) and 20% held out. The training fit is $\log Q \sim \log p + \text{sale flag} + \text{month-of-year} + \text{Jewish-holiday flags} \mid \text{product} \times \text{store}$, with no price-ending terms; the sale flag marks weeks priced at least 3% below the neighboring weeks on both sides. I then compute residuals—actual minus predicted log quantity—on the 2013 hold-out (13,746 observations) and on all of 2014–15 (151,807 observations). Because the model already conditions on price, sales, seasonality, and product-store identity, the residuals are compared as simple means by price-ending class, with no further controls, clustering by store.

Figure A-10 plots monthly mean residuals, split by ending class in panel (a) and pooled in panel (b). The pooled series shows that the model projects forward well: hold-out months average $+0.01$, and the 2014–15 months drift to -0.05 by 2015. The by-ending series shows the transition. In 2013, prices ending in .00–.09 are relatively rare (11% of observations), so the class mean is imprecise; demand at .00 was largely unknown, in the data and to the firms. In the first half of 2014 the .00–.09 mean drops to 0.14 log points below prediction, and it is estimated precisely, exactly because of the mass transition to .00: for many product-stores the move was a bad idea (quantities came in below what their own 2013 history predicted), and the firms had not yet realized it. Relative to the .90–.99 class, the residual gap is $+0.03$ (SE 0.02) on the hold-out, $+0.14$ (0.02) in 2014H1, -0.04 (0.02) in 2014H2, and -0.07 (0.02) in 2015. As the correction proceeds, the product-stores still at .00 become an increasingly selected set: the surviving .00s are the ones that were not underperforming. The standard errors widen again, and the later class means measure survivorship rather than the penalty.

Two notes on inference and robustness. Standard errors cluster by store and treat the residuals as data. A store-block bootstrap of the full pipeline—redrawing stores with replacement, re-splitting

2013, retraining, and recomputing the class means, 200 replications—gives standard deviations of 0.022 for the hold-out gap and 0.019 for the 2014H1 gap, against analytic standard errors of 0.022 and 0.020: the training-stage estimation error adds essentially nothing. (The product-store fixed-effect noise is within store, so the clustering absorbs it; the delta-method contribution of the trained slope coefficients is 0.003–0.005.) Across replications the hold-out gap averages +0.005, so the +0.03 point estimate partly reflects the particular hold-out draw; the 2014H1 bootstrap interval is [+0.10, +0.17]. The gap is also not an artifact of price dynamics: replacing the sale flag with same-price spell-length bins (1, 2, 3, 4+ weeks) interacted with the direction of the spell’s price relative to the previous spell gives a 2014H1 gap of +0.10 (SE 0.02) with a hold-out benchmark of −0.01 (on the unfiltered panel, which the one-week bins require). The analogous substitution in the specification of Table 1, putting direction-by-length spell types in place of the spell-length-decile and sale-flag fixed effects, moves its endings-IV gap from +0.10 to +0.11.

Figure A-10: 2013-trained out-of-sample monthly residuals



residuals (actual minus predicted log quantity) from a demand model trained on 80% of 2013: $\log Q$ on $\log p$, a sale flag, month-of-year and Jewish-holiday controls, and product $\times$ store fixed effects, with no price-ending terms. The shaded region is the 2013 20% hold-out; 2014–15 is fully out of sample. Panel (a) splits observations by price-ending class; panel (b) pools all endings. Bands are 95% confidence intervals, clustered by store.

K Heterogeneity in θ

One concern is whether the left-digit-bias estimate masks heterogeneity in θ across products and price levels. Three analyses address this.

K.1 Cross-product heterogeneity

Running the demand specification separately on each of the 15 (of 17) StoreNext products with sufficient sample size gives gaps ranging from -0.16 to $+0.29$, with about half the estimates not statistically different from zero individually. Seventeen products is too few for reliable per-product estimation; the pooled estimate is the right summary.

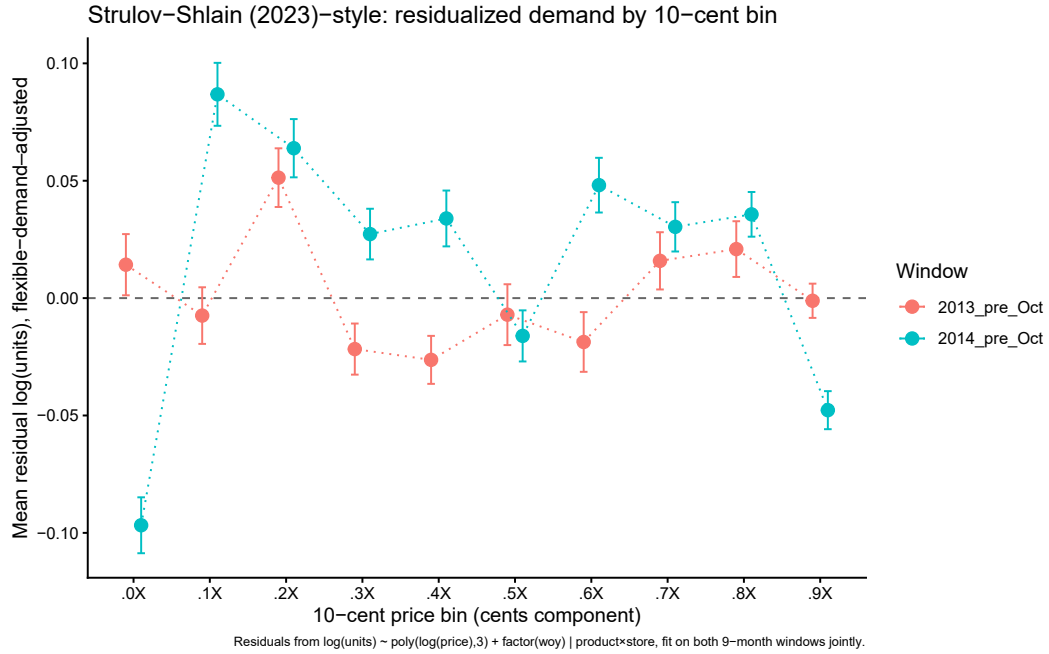
K.2 Variation with price level

Splitting product-stores into tertiles by within-product price level and re-estimating the demand specification with tertile-interacted ending dummies, the gap is $+0.047$ in the low tertile, $+0.034$ in the middle, and -0.025 in the high tertile. Pairwise tests: low vs. mid $p = 0.77$; mid vs. high $p = 0.04$; low vs. high $p = 0.13$. The high tertile gap is significantly lower than the middle, opposite to the prediction that left-digit bias is stronger at higher price levels.

K.3 Strulov-Shlain (2023)-style binned residuals

A different visualization: residualize $\log Q$ against a flexible price function + week-of-year + product $\times$ store + window fixed effects on a joint sample of 2013-pre-October and 2014-pre-October, then plot mean residuals by 10-cent bin within each window. The .0X bin moves from a slightly positive residual in 2013-pre-October ($+0.014$) to a strongly negative residual in 2014-pre-October (-0.097); the .9X bin moves from essentially zero (-0.001) to negative (-0.048); the .5X bin moves from -0.007 to -0.016 . The other seven bins do not shift comparably (Figure A-11). The dimes whose composition flipped between the two windows (.0X, .9X, .5X) are the ones that show the post-reform demand discontinuity, which is what the model predicts.

Figure A-11: Strulov-Shlain (2023)-style binned residuals, 2013 pre-Oct vs 2014 pre-Oct



Mean residuals from a flexible-demand regression by 10-cent price bin, separately for 2013 pre-October and 2014 pre-October. The .0X, .9X, and .5X bins are exactly the ones whose composition flipped between the two windows.

L Alternative accounts of the .00 mass: details

Table A4 compares the candidate accounts of the transition to .00. It sat at the head of Section 4 in the submitted version and is reproduced here in full, with a verbal summary retained in the main text.

Deliberate experimentation: the full inventory. Section 4.3 summarizes four features of the data that are hard to reconcile with a centrally designed experiment. First, the assignment does not look designed: there is limited movement from .90 to .00 overall, and within a product-store there is almost no back-and-forth between endings, with .90 absorbing once tried. Second, the assignment has little additive chain structure yet is correlated across chains: product identity carries the bulk of the explained variance in which product-stores moved to .00, so competing chains moved to .00 for the same products. Independent experimenters choosing their own test sets would not necessarily pick the same products. Third, information revealed in one product-store did not translate quickly throughout the chain. The cross-store contrast needed to evaluate the .00 shortfall existed within weeks of the reform (the variation the shortfall measure of Section 3.5 uses), yet chain-wide reversion took months and tracked the precision of the firm’s own benchmark. The spillover decomposition shows switches clustering in same-week batches within a store, though not covering

Table A4: Alternative accounts of the transition to .00, and how each fits the evidence

Account	What it would predict	What the data show	Where
Full-information optimization	No .00 endings at any point; an immediate switch to .90	20% of prices ended in .00 at the onset	§3
Stable but wrong parameter beliefs (bias underestimated)	One stable parameter belief fits pricing before and after the reform; still no .00 if it rationalizes decades of .99-use	Within the specified demand class, perceived $\hat{\theta} = 0.02-0.04$ lies inside the no-.00 region; one stable parameter does not produce both the .90 bunching and the .00 mass	§4.2
Structured optimization under parameter uncertainty	Avoid .00 from the start within the specified demand class; tend to correct in order of forgone profit; a shared-parameter update carries across related products	The .00 mass at the onset; recovery order follows pre-reform .99-reliance more closely than θ_0 ; the order products adopted .00 is unrelated to the order they left it ($\rho = 0.07$)	§4.1
Non-monotonic demand with ending-specific intercepts	.00 can be subjectively optimal if it has a favorable intercept; updating each ending only after it is tried can generate product-by-product reversion	Can rationalize the path, but only by leaving the ranking of endings essentially unrestricted; with action-specific values learned from own outcomes, the account has little ex ante empirical bite and is behaviorally close to model-free learning	§4.2
Chain-wide belief heterogeneity	A chain-wide propensity organizes which products move to .00	Additive chain identity explains about a tenth of the predictable variation, although chains implement product-specific choices consistently across stores	§3.3
Deliberate experimentation	A designed assignment; relatively quick use of in-house information; repeated testing	Little back-and-forth within a product; competing chains often move the same products to .00; months to act on comparisons available within weeks; little repeated testing	§4.3
Wholesale-cost shift	The .00 mass mirrors the cost distribution	At most 0.25% .00 under the simulated cost levels and grains; the mass unwinds without an observed corresponding cost change	App. L
Manufacturer-dictated endings	.00 imposed from upstream, and unlikely to unwind product by product on its own	The sampled products have no observed vertical price restrictions; the .00 mass unwinds within a year	App. L
Consumer placation / price image / reputation concerns	Consumers favor .00; adoption spans a chain's assortment	.00 is associated with substantially lower demand under the ending IVs; adoption has little chain-wide structure and is abandoned product by product	App. L
Confusion about competitors, not demand	Corrections move with rivals' endings	Recovery is predicted by each product-store's demand shortfall and the product's prior .99-reliance	App. L
Decision or attention costs	Adjustment costs are important; choices may cluster at the relevant decision node; a potentially wide good-enough band	Repricing was common; no chain-wide or store-wide node concentrates the initial .00 choices; reversion is predicted by shortfalls	§4.4
Random implementation errors	Approximately symmetric, stationary, scattered errors	A one-sided 20% mass at the reform, a largely one-way flow back to .90, and a correction predicted by demand feedback	App. L
<i>Rules of thumb audited against accessible benchmarks</i>	An idiosyncratic onset; a correction paced by what each firm can detect; local learning; stopping once good enough	Fits the combined patterns without assigning unrestricted latent values to each untried ending	§4.1

Each row states an alternative account of the post-reform transition, the pattern that would distinguish it, and how it fits the evidence. The data do not logically exclude every flexible alternative, but the rows are not equally plausible: the last row is the paper's favored account because it explains the combined patterns with fewer auxiliary assumptions. Rows marked "App." are detailed in Appendix L; the rest are discussed in the indicated subsection.

all of a store’s switches, with no detectable gradual diffusion across a chain’s stores (Appendix G). Fourth, exploration was rarely repeated: once .90 cleared the benchmark, firms did not continue to test alternatives. A belief shock at headquarters would instead organize exploration at the chain level.

Structured versus unstructured learning. A classic bandit treats arms as independent: what a firm learns about .00 tells it nothing about .90 until .90 is tried. That is the absence of a model generalizing across actions, and it is the opposite of a demand model, which links all endings, so a firm holding one would not try a dominated .00 in the first place. The operations-research literature on pricing while learning makes the same point: a model-based learner experiments only among candidate optima, and its worst case is converging to an *uninformative* price, never a *dominated* one (Harrison et al., 2012). A Bayesian bandit or Thompson sampling, which samples an arm in proportion to the posterior probability that it is best (Russo and Van Roy, 2014), moves mass onto .00 only if firms are genuinely unsure of its value. Whether information about one action updates the value of others is what separates structured from unstructured learning (Schulz et al., 2020), and even the mild structure that demand slopes downward lets a tried price inform untried ones (Misra et al., 2019).

Decision costs: the organizational evidence. The initial .00 choices do not cluster in a single organizational node (Section 3.4), which makes a single chain-wide or store-wide attention bottleneck harder to sustain while leaving product-specific evaluation costs available. Attention was also necessarily devoted to the transition, since firms had to change almost all prices, and a firm not attending to differences between products could have priced everything at .90, which is a cheap rule. Chains shifted away from .00 at particular moments (Figure A-3), and that discreteness is hard to reconcile with menu costs or other adjustment frictions, which would smooth the decline. Choosing .00 also cost more downstream: that path incurred a second round of price adjustment plus a demand penalty of at least 6.5% on affected products.⁴²

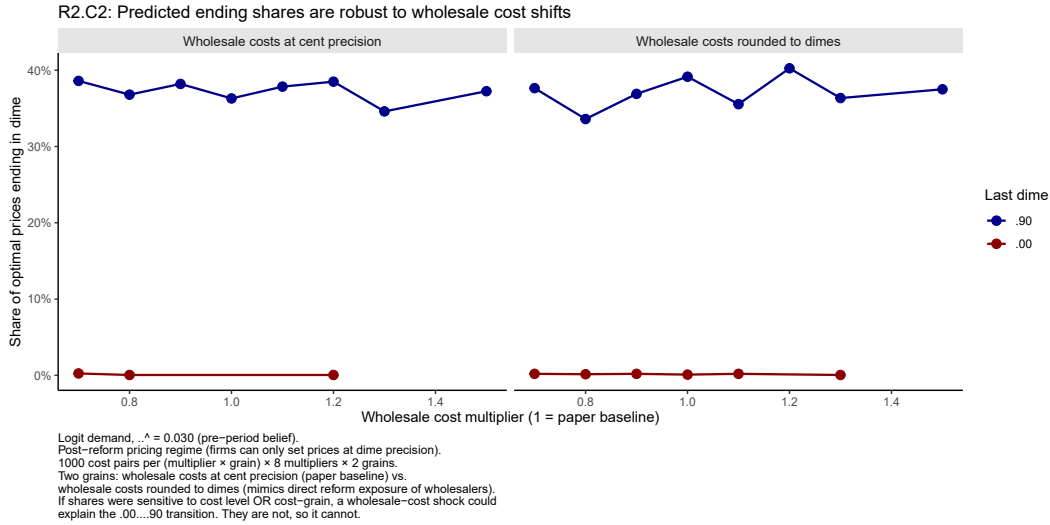
The rest of this appendix gives the evidence behind the rows of Table A4 that Section 4 eliminates without discussion.

Wholesale-cost shift. One possibility is that the reform indirectly shifted wholesale costs (or rounded wholesale prices to dimes), so that the .00 mass reflects the cost distribution rather than a heuristic mistake. Within the maintained demand model, .00 remains dominated across the cost distributions examined. I simulate $N = 1,000$ cost draws from the baseline Weibull(4, 6) distribution per cell and compute the optimal post-reform price under the pre-reform perceived bias $\hat{\theta} \approx 0.03$ and baseline logit demand, varying both the cost level and grain. Across all sixteen cells, the .00

⁴²The pattern is analogous to Ugandan firms that sometimes misreport taxes in the direction that raises their own liability (Almunia et al., 2024): a costly-action account must explain why the error favors a readily avoidable alternative.

share of optimal prices is in $[0\%, 0.25\%]$ and the .90 share is in $[33.6\%, 40.2\%]$, with no monotone dependence on cost level or grain (Figure A-12). Thus the simulated cost shifts do not reproduce the observed 20% .00 mass.

Figure A-12: Optimal-price ending shares as a function of wholesale-cost level and grain



The figure shows the share of optimal post-reform prices ending in each dime, across 16 (cost-multiplier $\times$ cost-grain) cells with $N = 1,000$ cost draws each. Cost multiplier varies between 0.7 and 1.5; cost grain is either cents or dimes-rounded. The .00 share remains near zero regardless.

Manufacturer-dictated endings. A second possibility is that manufacturers dictated .00 endings through vertical pricing agreements. Three facts rule this out. The products in the sample are not subject to vertical price restrictions.⁴³ Prices and price endings vary from period to period, even across stores of the same chain (Figure A-8). And a manufacturer-imposed .00 ending would not unwind on its own, whereas the .00 mass is undone within a year, driven by retailers.

Consumer placation and price image. A third possibility is that firms chose .00 endings to soothe consumers during a confusing policy change. The demand estimates say the opposite: under the strongest IV specifications α^{00} is -0.1 to -0.16 , so consumers respond to .00 far more negatively than to .90, and a firm choosing .00 imposed a real demand penalty on itself rather than soothing anyone. Placation would predict the reverse: firms moving toward the endings consumers prefer.⁴⁴ A portfolio version—chains adopting round prices across products to signal simplicity or transparency—fails on the same evidence, plus one more piece. A price-image strategy is a chain-level portfolio choice, so it would cluster .00 adoptions by chain across each chain's assortment,

⁴³No private-label products are included in the StoreNext sample; manufacturer-suggested retail prices are legally scrutinized in Israel.

⁴⁴If anything, the news article in footnote 14 presents a change from 5.99 to 6.00 as an example of firms using the reform to raise prices.

whereas the transitions to .00 are organized between products, with endings varying product by product within the same store (Section 3.4). An image worth a 6.5% demand penalty per flagged product would also not be abandoned product by product, paced by each product’s own shortfall, while the consumer-protection context was unchanged. Further, if firms were steering by press risk, the coverage told them to avoid round-ups: the December 2013 article presents 5.99 becoming 6.00 as an example of firms using the reform to raise prices (footnote 14). A fifth of prices moved to .00 anyway.

Confusion about competitors, not demand. A fourth possibility is strategic: firms chose .00 anticipating that competitors would too, then unwound when they did not. Three facts push against it. Chains moved similarly to one another early on, with only modest between-chain heterogeneity in the .00 spike and in the speed of recovery (Figure A-3). Each product’s recovery is paced *within* its chain by that product’s own pre-reform heuristic strength and its own realized demand shortfall (Section 3.5), not by rivals’ endings; a strategic unwinding would instead have each product’s correction co-move with competitors’ moves. And the environment was otherwise stable (no entry, no exit, unchanged assortments), so the natural reading is confusion about own demand rather than about strategic interaction.

Random implementation errors. A fifth possibility is mechanical: firms optimize but implement prices with occasional error, as a trembling hand. A simple symmetric, stationary error process fits the observed pattern poorly: the data show a 20% mass concentrated on one ending at the reform, a largely one-way flow from .00 back to .90, and a correction predicted by demand feedback and prior .99-adherence. A fixed-size tremble on the coarser post-reform grid would also tend to raise post-reform forgone profits, whereas the calibrated series ends below its pre-reform level. More structured or state-dependent error processes are possible, but require additional assumptions to match these asymmetries.

M Chain heterogeneity and a volatility-learning robustness check

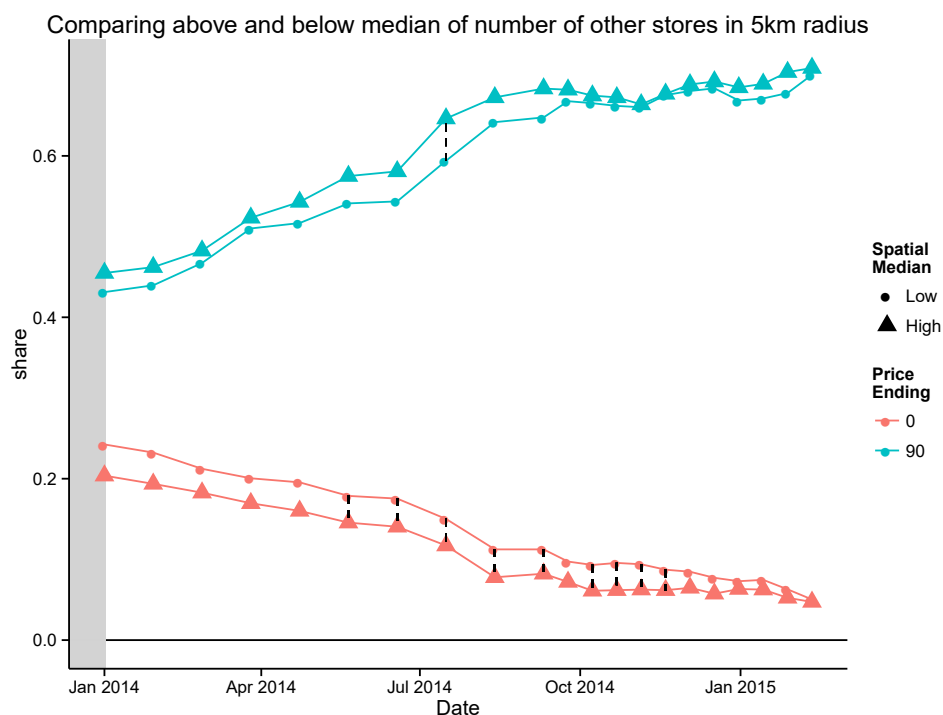
M.1 Chain heterogeneity in convergence speed

Across the seven chains that converge and have at least three stores, regressing weeks-to-convergence on chain characteristics gives the following. Alone, $\log(n_{\text{stores chain}})$ has coefficient +21.8 (SE 5.3, exact $p = 0.009$; restricted wild-bootstrap $p = 0.008$). In the combined model with positioning and $\log(\text{average price})$, the coefficient is +21.5 (SE 4.7, $p = 0.019$ at three residual degrees of freedom; wild-bootstrap $p = 0.077$); the discount-vs-premium positioning dummy is not significant ($p = 0.36$), nor is $\log(\text{average price})$ ($p = 0.88$). The pattern is suggestive: bigger chains

take longer to converge. With seven chains, this should be read as suggestive rather than as a sharp identifying test.

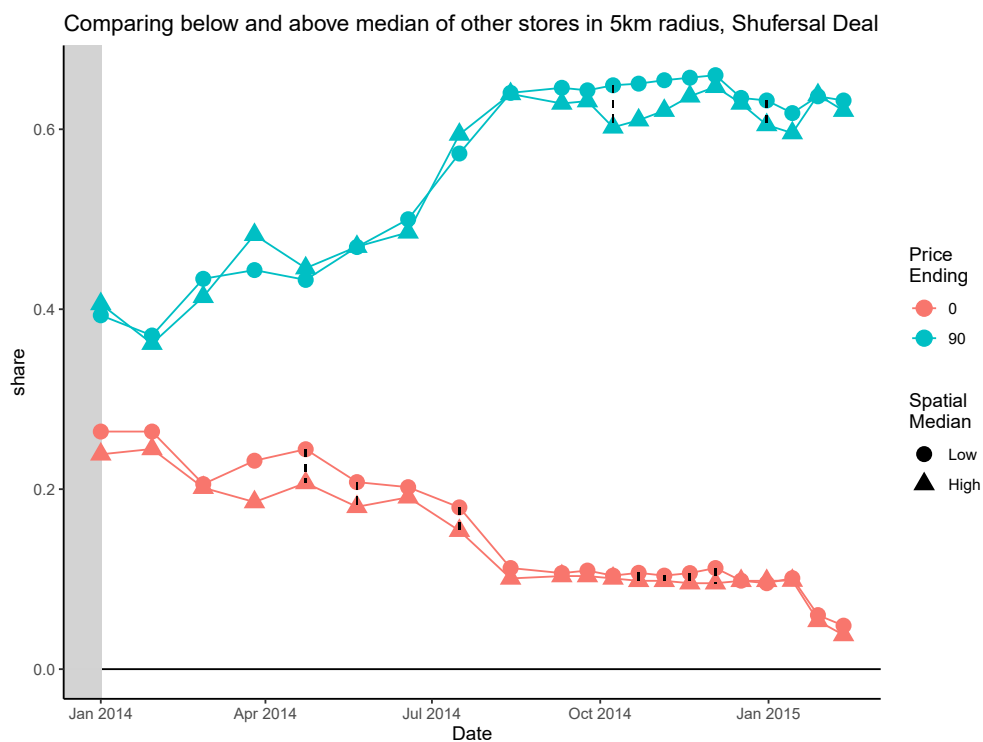
Figure A-13: Spatial competition and shares of price ending

(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

The figure shows the



shares of .00 and .90 prices at the post-period for stores that are above and below the median of number of local competitors.

The top panel shows all stores, and the bottom panel only the stores of the Shufersal Deal chain.

M.2 Pre-reform price-change frequency and learning speed

At the ICC *product* level, pre-reform price-change frequency correlates with faster convergence: for each product I compute the pre-reform price-change frequency (share of biweek-pairs with a price change) and the post-reform time-to-50%-.90-share, giving a Spearman correlation of $\rho = -0.57$ ($N = 39$ product-level observations). This product-level pattern does *not*, however, survive at the product-store level under the congruent first-.90 outcome used in Section 3.5: measuring each product-store’s recovery by its own first-.90 date rather than by a 39-product censored “weeks-to-50%-.90” aggregate, the relationship is null, and the multivariate recovery-order regression in Section 3.5 assigns pre-reform change frequency no role (+0.3 weeks per standard deviation, $p = 0.83$). I therefore do not treat volatility-driven learning speed as an established finding; the robust product-level predictor of recovery order is pre-reform .99-adherence.

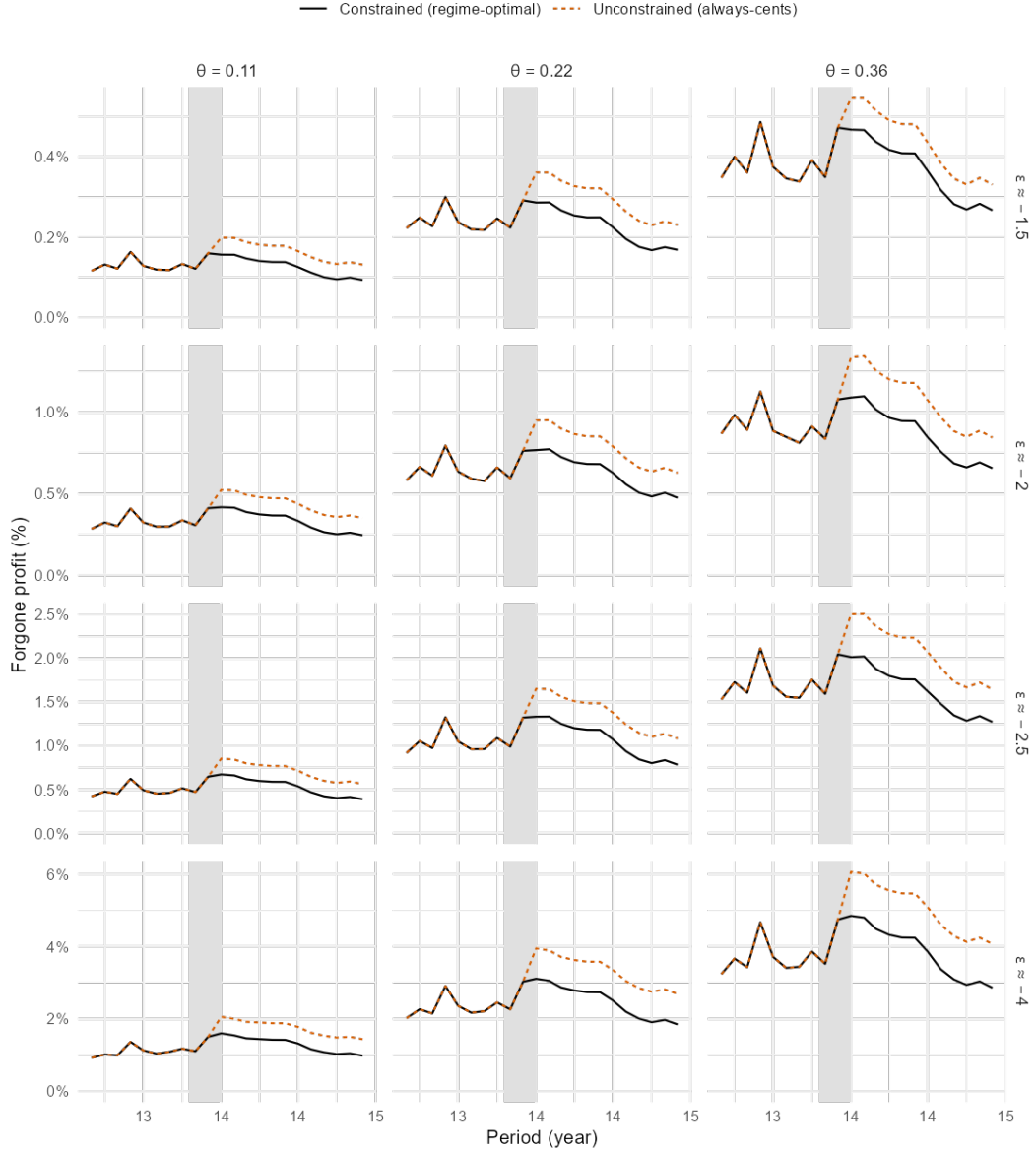
N June-2013 ksi spike diagnostic and smoothing variants

N.1 The June-2013 ksi spike is a model-flexibility artifact

The bootstrap supply-side parameter time series shows a ksi spike from ~ 0 in April 2013 to ~ 0.165 in June 2013, while the aggregate price-ending shares are stable over the same period. The MSM input is dime-binned; at the dime level, June 2013 has a 3–4 percentage-point shift from the .9X dime to the .0X and .5X dimes, which the MSM absorbs as a ksi spike even though the cent-level structure is unchanged. A stripped-down rule-of-thumb model fit to (a) April 2013 actual, (b) April + 3pp synthetic shift to .00, (c) June 2013 actual gives $\text{ksi} \in \{0.006, 0.092, 0.134\}$, matching the cluster’s 0.165 estimate almost exactly. The spike is a model-flexibility artifact, not a real signal about firm rounding behavior.

N.2 Robustness across the (θ, ε) grid

Figure A-14 shows the forgone-profit time series across a 3×4 grid of (θ, ε) values, under the regime-constrained benchmark (headline metric) and the always-cents-unconstrained benchmark. The elasticity columns are calibrated so the simulated *mean* elasticity across the cost draws equals -1.5 , -2.0 , -2.5 , and -4.0 ; the whole grid uses the same unweighted aggregation, so the $(\theta = 0.22, \varepsilon = -2.0)$ constrained cell reproduces Figure 5 exactly. The post-reform decline below the pre-reform level occurs in every cell.

Figure A-14: Forgone profits across the (θ, ε) grid, constrained and unconstrained benchmarks

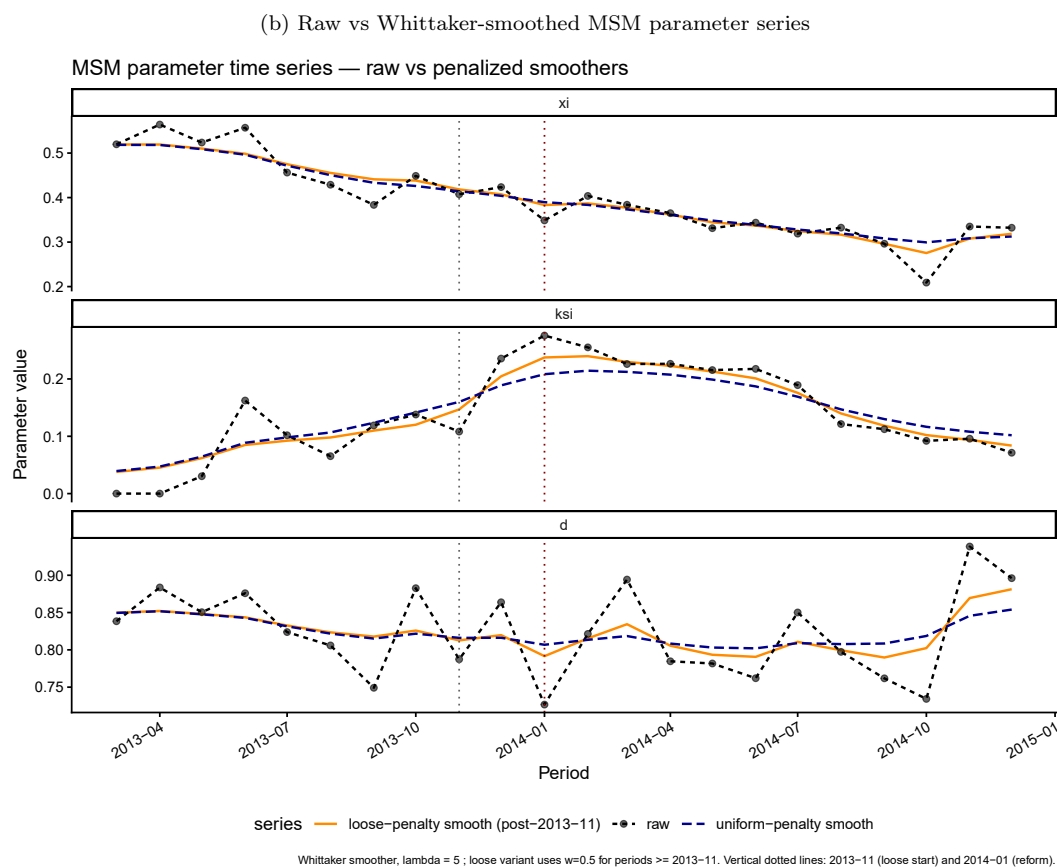
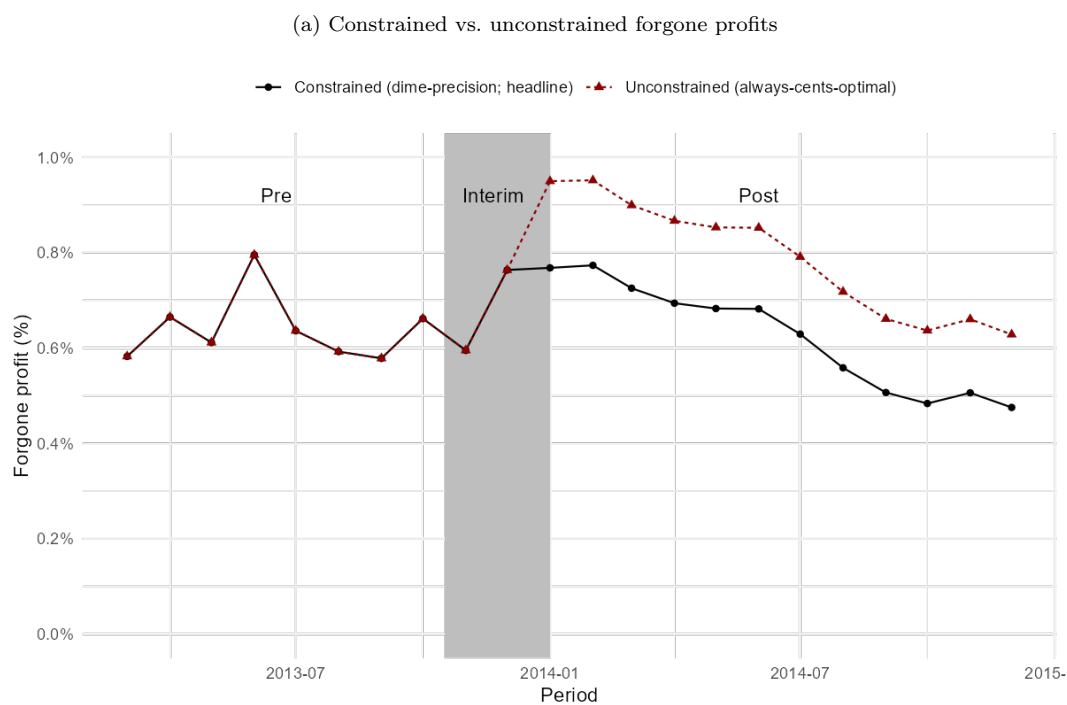
The twelve cells vary the demand parameters: rows = mean elasticity $\varepsilon \in \{-1.5, -2.0, -2.5, -4.0\}$ (the logit own-price coefficient is calibrated per row so the mean of $-\alpha p(1-s)$ across draws hits each target), columns = $\theta \in \{0.11, 0.22, 0.36\}$. Each elasticity row has its own vertical scale. Within each cell, the solid line is forgone profit against the regime-constrained optimum (the headline metric) and the dashed line against the always-cents-unconstrained optimum. Forgone profit is averaged unweighted across draws, so the $(\varepsilon = -2.0, \theta = 0.22)$ solid line is Figure 5.

N.3 Whittaker smoother on the (xi, ksi, d) MSM time series

A post-hoc first-difference-penalized smoother (Whittaker, $\lambda = 5$) on the monthly median MSM parameter series absorbs roughly 78% of the June-minus-April 2013 ksi spike ($+0.118 \rightarrow +0.027$) under a uniform penalty, and a comparable share under a looser-post-2013-11 penalty. The

smoothed parameter series, when fed back through the counterfactual pipeline, produces a forgone-profit time series with essentially the same pre- and post-reform pattern as the raw-median input. The constrained vs. unconstrained benchmark comparison (Figure A-15a, at $\varepsilon = -2.0$ calibration, shown alongside Figure A-15b) places the cost of the dime-precision constraint alone at about 0.15 percentage points throughout the post-reform period.

Figure A-15: Counterfactual smoothing variants

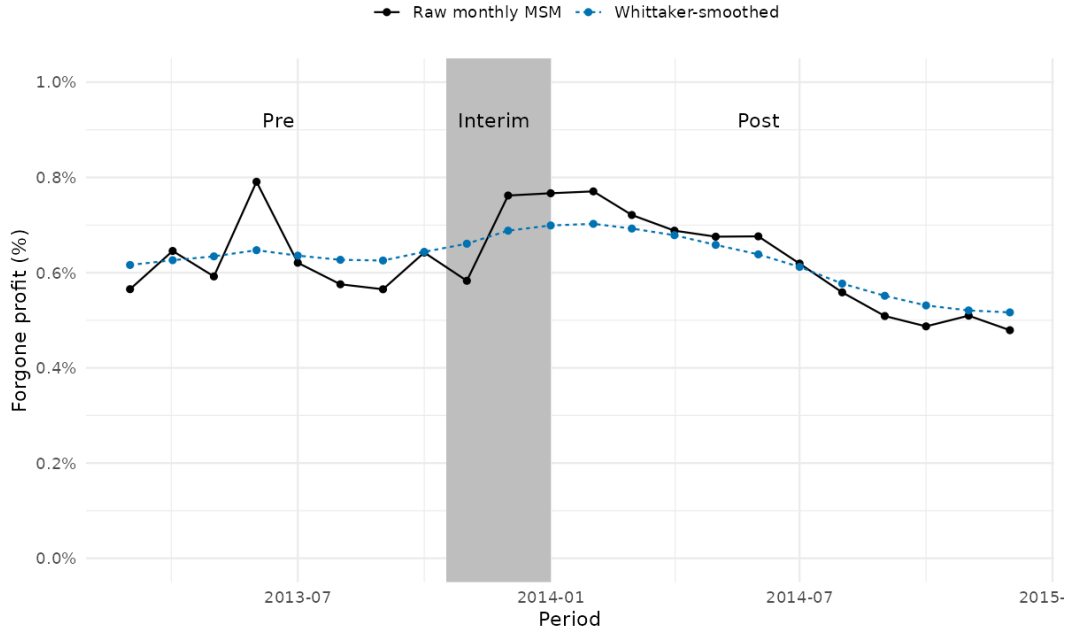


Top panel: constrained (dime-precision) vs. unconstrained (always-cents-optimal) forgone profits at the $\varepsilon = -2.0$ calibration.

The vertical gap is the cost of the dime constraint, about 0.15 pp. Bottom panel: monthly median MSM estimates of (ξ_t, κ_t, d_t) in black points; uniform-penalty Whittaker smoother in dark blue; looser-post-2013-11 variant in dark orange.

Vertical dotted lines mark 2013-11 (the loose-penalty boundary) and 2014-01 (the reform).

Figure A-16: Counterfactual forgone profits with smoothed MSM parameters



Each curve is the monthly forgone-profit estimate from the structural counterfactual pipeline, using a different MSM parameter input: raw per-month medians (black), uniform-penalty smoothed (blue), and looser-post-2013-11 smoothed (orange). The pre-/post-reform pattern is essentially unchanged.

O Cleaning sensitivity

Re-running the chain-interacted Hausman IV specification at five cleaning thresholds gives a monotone pattern: stricter cleaning yields a larger gap.

Table A5: Cleaning sensitivity: gap by cleaning threshold

Cleaning rule	90-vs-00 gap (SE)	[p-value]	<i>N</i>	share kept	Each
Strict (spell ≥ 4 weeks)	+9.5% (3.1)	[0.003]	149k	27%	
Strict (admissible-price only)	+9.0% (2.3)	[<0.001]	180k	33%	
Default (spell > 2 or admissible)	+7.0% (1.8)	[<0.001]	247k	45%	
Loose (spell ≥ 2 or admissible)	+6.1% (1.5)	[<0.001]	294k	54%	
Loosest (no filter)	+2.1% (0.9)	[0.025]	524k	96%	

row reports the gap $\alpha^{90} - \alpha^{00}$ from the chain-interacted Hausman IV specification (columns 1–6 of Table 1), run on samples cleaned at five different thresholds; store-clustered delta-method standard errors in parentheses (in percentage points) and exact *p*-values in brackets. The smaller gaps in looser samples are consistent with attenuation from short-spell and ambiguous-price observations, although sample composition also changes.

The default cleaning rule sits in the middle of this monotone pattern. The decline in the gap as short-spell and ambiguous-price observations are admitted is consistent with attenuation, although sample composition also changes. I therefore use the default estimate as a cautious estimate rather than treating the sensitivity exercise as proof of the direction of bias.

P Forgone Profits Computation

This appendix provides details on the computation of forgone profits reported in Section 3.

P.1 Demand Model

I model demand using a two-product logit with left-digit bias. Consumer j chooses product $i \in \{1, 2, \text{outside}\}$ to maximize:

$$u_{ij} = \delta_i - \alpha \hat{p}_i + \varepsilon_{ij} \quad (4)$$

where δ_i is the product intercept, $\alpha > 0$ is price sensitivity, $\hat{p}_i = (1 - \theta)p_i + \theta(\lfloor p_i \rfloor + \Delta)$ is the perceived price incorporating left-digit bias, and ε_{ij} is Type-I extreme value. The outside option has $\delta_0 = 0$ and $p_0 = 0$.

Market shares follow the standard logit form:

$$s_i(p_1, p_2; \theta) = \frac{\exp(\delta_i - \alpha \hat{p}_i)}{1 + \exp(\delta_1 - \alpha \hat{p}_1) + \exp(\delta_2 - \alpha \hat{p}_2)} \quad (5)$$

I use logit rather than constant-elasticity (log-log) demand because it produces more conservative estimates of forgone profits. Under log-log demand, losses from suboptimal pricing are larger; qualitative patterns are similar.

P.2 Calibration

Parameters are calibrated as follows:

- **Product intercepts:** $\delta_1 = 1$, $\delta_2 = 0.5$. These are normalized; the key is that products differ in baseline appeal.
- **Price sensitivity:** $\alpha = 0.18$, set so the simulated *mean* own-price elasticity across the cost draws—the average of $-\alpha p(1-s)$ —equals the estimated -2.0 . Forgone profit is then averaged unweighted across the draws. (An earlier version reported a weighted average with an $(\varepsilon+2)^{-2}$ kernel; because that kernel is near-singular at -2.0 it made the reported level depend on a thin subset of draws, so the current calibration instead centers the draw distribution itself at

−2.0 and averages plainly. At $\alpha = 0.18$ the mean, median, and kernel-weighted forgone-profit series coincide.) The period-interacted IV estimates in Appendix J put the elasticity at −2.01; the robustness grid in Appendix N.2 spans −1.5 to −4.

- **Left-digit bias:** $\theta = 0.22$, just below the demand-side range of 0.25–0.38 estimated in Section 3, so the calibration is conservative on this margin.
- **Focal ending:** $\Delta = 0$.
- **Cost distribution:** Costs are drawn from a Weibull distribution with shape parameter $k = 4$ and scale parameter $\lambda = 6$. This distribution is chosen so that the resulting price distribution (under optimal pricing) roughly matches the observed price range and distribution shape.

P.3 Simulation Procedure

Cost draws. I draw 1,000 costs for product 1 from the Weibull distribution. For each product-1 cost, I independently draw 50 costs for product 2, yielding 50,000 $(\text{cost}_1, \text{cost}_2)$ pairs.

Optimal pricing. For each cost pair, I solve for the prices (p_1^*, p_2^*) that maximize total profits given both prices:

$$(p_1^*, p_2^*) = \arg \max_{p_i, p_j} s_i(p_i, p_j; \theta) \cdot (p_i - c_i) + s_j(p_i, p_j; \theta) \cdot (p_j - c_j) \quad (6)$$

The solution is computed twice for each cost pair:

1. **Pre-reform:** prices can be set to the cent (e.g., 2.99, 3.00, 3.01, ...)
2. **Post-reform:** prices restricted to dimes (e.g., 2.90, 3.00, 3.10, ...)

Profits under optimal pricing. For each cost draw i :

$$\Pi_i^* = s_1(p_1^*, p_2^*; \theta) \cdot (p_1^* - c_{1i}) + s_2(p_1^*, p_2^*; \theta) \cdot (p_2^* - c_{2i}) \quad (7)$$

P.4 Rule-of-Thumb Heuristic

The descriptive model of actual pricing is a “rule-of-thumb” heuristic with three parameters estimated each month: $\omega_t = (\omega_t^{99/90}, \omega_t^{00}, r_t)$.

Step 1: Naive price. Given cost c and price elasticity ε (computed from the logit model at average prices), compute the naive optimal price ignoring left-digit bias:

$$p^{\text{naive}} = \frac{c \cdot \varepsilon}{1 + \varepsilon} \quad (8)$$

This is Lerner’s rule for a firm facing constant-elasticity demand.

Step 2: Rounding decision. Let $p^{\text{round}} = \text{round}(p^{\text{naive}})$ be the nearest integer to the naive price (higher or lower). If $|p^{\text{naive}} - p^{\text{round}}| \leq r_t$ (the naive price is “close” to a round number):

- With probability $\omega_t^{99/90}$: set price to $p^{\text{round}} - 0.01$ (pre-reform) or $p^{\text{round}} - 0.10$ (post-reform)
- With probability ω_t^{00} : set price to p^{round}
- With probability $1 - \omega_t^{99/90} - \omega_t^{00}$: keep p^{naive}

If the naive price is not close to a round number, keep p^{naive} .

Estimating ω_t . For each month t , I estimate $(\omega_t^{99/90}, \omega_t^{00}, r_t)$ by matching the model-implied price distribution to the observed distribution. The procedure is:

1. Take the empirical shares of .99/.90s and .00s in month t
2. For candidate parameter values, simulate the heuristic pricing for each cost draw
3. Compute the implied shares of each price ending
4. Choose parameters to minimize the distance between implied and empirical shares

This procedure assumes the cost distribution is stable over time (estimated once from pre-reform data using the method in Appendix Q) while the heuristic parameters ω_t vary month-by-month.

P.5 Computing Forgone Profits

For each month t and cost draw i :

1. Compute optimal prices (p_{1i}^*, p_{2i}^*) and optimal profit Π_i^*
2. Apply the heuristic with estimated $\hat{\omega}_t$ to get rule-of-thumb prices (p_{1i}^r, p_{2i}^r)
3. Compute rule-of-thumb profit: $\Pi_i^r = s_1(p_{1i}^r, p_{2i}^r; \theta) \cdot (p_{1i}^r - c_{1i}) + s_2(\cdot) \cdot (p_{2i}^r - c_{2i})$

Forgone profits in month t are the unweighted mean, across the cost draws, of the per-draw profit shortfall:

$$\text{Forgone}_t = \frac{1}{N} \sum_i \left(1 - \frac{\Pi_i^r}{\Pi_i^*} \right) \quad (9)$$

This measures the percentage of potential profits lost due to heuristic (rather than optimal) pricing. The own-price coefficient α is set (Appendix P, Calibration) so that the mean own-price elasticity across the draws equals the estimated -2.0 ; at that calibration the unweighted mean above coincides with its median and with an elasticity-weighted mean, so the reported level does not depend on the aggregation choice.

Q Perceived Left-Digit Bias Estimation

This appendix provides a detailed description of the Method of Simulated Moments (MSM) procedure used to estimate the “perceived” left-digit bias parameter θ implied by observed pricing behavior. The goal is to find the value of θ that, if firms were optimizing against left-digit biased demand with that belief, would best rationalize the observed distribution of prices. This procedure follows Strulov-Shlain (2023).

Q.1 Setup and Notation

Model. Firms face constant-elasticity demand with left-digit bias. The perceived price is:

$$\hat{p}(p; \theta) = (1 - \theta)p + \theta(\lfloor p \rfloor + \Delta) \quad (10)$$

where p is the actual price, $\lfloor p \rfloor$ is the floor (largest integer less than or equal to p), Δ is the focal price ending (set to 0 throughout), and $\theta \in [0, 1]$ is the left-digit bias parameter. Demand is $D(p; \theta) = A\hat{p}(p; \theta)^\epsilon$ where $\epsilon < -1$ is the price elasticity.

Firm optimization. A profit-maximizing firm with marginal cost c solves:

$$\max_p \Pi(p; \theta, c) = D(p; \theta) \cdot (p - c) \quad (11)$$

The first-order condition for an interior optimum (away from the discontinuity at round numbers) yields the relationship between optimal price and cost:

$$c(p; \theta, \epsilon) = p + \frac{\hat{p}(p; \theta)}{\epsilon(1 - \theta)} = p \frac{1 + \epsilon}{\epsilon} + \frac{\theta}{1 - \theta} \frac{\lfloor p \rfloor + \Delta}{\epsilon} \quad (12)$$

This equation maps each non-just-below price p to the unique cost c for which p is optimal (given θ and ϵ).

Key implication: bunching and missing prices. Under left-digit bias ($\theta > 0$), firms should bunch at “just-below” prices (e.g., X.99) and avoid prices just above round numbers. The *Next-Lowest Price* P_q is the lowest price above a just-below price q that is optimal for some cost. For any just-below price q (e.g., 2.99), there exists a threshold $P_q > \lfloor q \rfloor + 1$ such that:

- Prices in the interval $(\lfloor q \rfloor + 1, P_q)$ are never optimal for any cost (these are “missing” prices)
- The just-below price q is optimal for costs in some interval $[\underline{c}_q, \bar{c}_q]$
- Prices at or above P_q become optimal for sufficiently high costs

The Next-Lowest Price P_q is found by solving for the price $P > \lfloor q \rfloor + 1$ such that there exists a cost c where the firm is indifferent between q and P , and P satisfies the first-order condition at that cost. This is given implicitly by Equation 15.

Q.2 Data and Moments

Price data. The estimation uses the empirical distribution of prices from the ICC dataset. Let $\{p_1, p_2, \dots, p_N\}$ denote the observed prices (one observation per product-store-period).

Price grid. Define a price grid $\mathcal{P} = \{\rho_1, \dots, \rho_K\}$ covering the relevant range. For example, $\mathcal{P} = \{2.00, 2.10, 2.20, \dots, 3.90, 4.00\}$ at the dime level.

Empirical moments. For each price $\rho_k \in \mathcal{P}$, compute the empirical share:

$$S_k^{\text{data}} = \frac{\sum_i \mathbb{1}\{p_i \in [\rho_k, \rho_{k+1})\}}{N} \quad (13)$$

These empirical shares are the moments to be matched.

Q.3 Estimation Algorithm

The algorithm proceeds in two stages: (1) estimate the shape of the underlying price/cost distribution using prices unaffected by bunching, and (2) use the model to predict price shares as a function of (θ, ϵ) and find the parameters that best match the data.

Q.3.1 Stage 1: Estimate the Price Distribution Shape

Identify unaffected prices. Under the model, prices ending in X.99 exhibit excess mass (bunching), and prices just above round numbers (e.g., X.00 to X.28) are missing. Prices in the middle of each dollar range (e.g., X.29 to X.89) should be unaffected by left-digit bias and reflect the underlying cost distribution.

Fit a smooth distribution. Using only the unaffected prices, fit a parametric distribution to the empirical CDF. I use a logistic polynomial:

$$\hat{F}_p(p) = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 p + \beta_2 p^2 + \dots))} \quad (14)$$

Specifically:

1. Select prices in the “unaffected” range: for each dollar interval $[n, n + 1)$, use prices from $n.29$ to $n.89$ (or a similar range that excludes bunching at X.99 and missing prices near X.00).
2. Compute the empirical CDF at these prices.
3. Regress the logit of the empirical CDF on a polynomial in price: $\text{logit}(\hat{F}_p^{\text{emp}}) = \beta_0 + \beta_1 p + \beta_2 p^2 + \dots$
4. Use the fitted coefficients to define $\hat{F}_p(p)$ for all prices in $\mathcal{P}$.

This gives a smooth estimate of what the price distribution *would* look like in the absence of left-digit bias.

Q.3.2 Stage 2: Predict Price Shares Given (θ, ϵ)

For each candidate parameter pair (θ, ϵ) :

Step 2.1: Identify just-below prices and compute Next-Lowest Prices. For each just-below price q in the data (e.g., $q \in \{1.99, 2.99, 3.99, \dots\}$):

1. Solve for the Next-Lowest Price P_q using the indifference condition. Specifically, find $P_q > \lfloor q \rfloor + 1$ such that:

$$P_q + \frac{\hat{p}(P_q; \theta)}{\epsilon(1 - \theta)} = \frac{\hat{p}(P_q; \theta)^\epsilon \cdot P_q - \hat{p}(q; \theta)^\epsilon \cdot q}{\hat{p}(P_q; \theta)^\epsilon - \hat{p}(q; \theta)^\epsilon} \quad (15)$$

Solve numerically (e.g., using bisection or Newton’s method).

2. Record P_q . For example, if $\theta = 0.03$ and $\epsilon = -3$, solving for $q = 2.99$ might yield $P_{2.99} = 3.27$.

Step 2.2: Construct the predicted price grid. The model predicts three types of prices:

1. **Just-below prices** $\{q_1, q_2, \dots\}$: these have excess mass due to bunching
2. **Missing prices**: prices in $(\lfloor q_i \rfloor + 1, P_{q_i})$ for each q_i should have zero mass
3. **Regular prices**: prices at or above P_{q_i} (up to the next just-below price) follow the smooth distribution

Construct the augmented price grid $\mathcal{P}'$ that includes:

- All just-below prices $\{q_i\}$
- All Next-Lowest Prices $\{P_{q_i}\}$
- All regular prices between P_{q_i} and q_{i+1}

Step 2.3: Map prices to costs. For each price $p \in \mathcal{P}'$, compute the cost $\underline{c}_p$ such that p is the profit-maximizing price:

$$\underline{c}_p = p \frac{1 + \epsilon}{\epsilon} + \frac{\theta}{1 - \theta} \frac{\lfloor p \rfloor + \Delta}{\epsilon} \quad (16)$$

Step 2.4: Recover the cost distribution. The cost distribution is derived from the price distribution using the change-of-variables implied by the first-order condition. Since $\hat{F}_p$ gives the CDF of prices (from Stage 1), and the FOC maps prices to costs, we have:

$$\hat{F}_c(c) = \hat{F}_p(p(c)) \quad (17)$$

where $p(c)$ inverts the FOC. Substituting:

$$\hat{F}_c(\underline{c}_p) = \hat{F}_p \left(\underline{c}_p \cdot \frac{\epsilon}{1 + \epsilon} - \frac{\theta}{1 - \theta} \frac{\lfloor p \rfloor + \Delta}{1 + \epsilon} \right) \quad (18)$$

Step 2.5: Compute predicted price shares. For each price p_i in the augmented grid $\mathcal{P}'$:

$$\hat{S}_{p_i}(\theta, \epsilon) = \hat{F}_c(\underline{c}_{p_{i+1}}) - \hat{F}_c(\underline{c}_{p_i}) \quad (19)$$

This gives the predicted share of prices at p_i : it equals the mass of costs for which p_i is optimal.

For **just-below prices** q : the predicted share includes all costs in the bunching region:

$$\hat{S}_q(\theta, \epsilon) = \hat{F}_c(\underline{c}_{P_q}) - \hat{F}_c(\underline{c}_q) \quad (20)$$

where $\underline{c}_q$ is the lowest cost for which q is optimal, and $\underline{c}_{P_q}$ corresponds to the Next-Lowest Price.

For **missing prices** (between $\lfloor q \rfloor + 1$ and P_q): set $\hat{S}_p = 0$.

Step 2.6: Aggregate to the empirical price grid. The augmented grid $\mathcal{P}'$ may have prices (like $P_q = 3.27$) that don't correspond to actual price points in the data. Aggregate predicted shares to match the empirical grid $\mathcal{P}$:

- Assign the share at P_q to the nearest price ending in 9 below it. For example, if $P_{2.99} = 3.27$, assign its share to 3.19 (or the next admissible 9-ending price).
- Sum shares for prices that map to the same grid point.

Q.3.3 Stage 3: Minimize Distance

Objective function. Choose (θ, ϵ) to minimize the sum of squared differences between predicted and empirical shares:

$$(\hat{\theta}, \hat{\epsilon}) = \arg \min_{\theta, \epsilon} \sum_{p \in \mathcal{P}^*} \left(\hat{S}_p(\theta, \epsilon) - S_p^{\text{data}} \right)^2 \quad (21)$$

where $\mathcal{P}^*$ excludes the prices used to fit the distribution in Stage 1 (to avoid using the same data twice).

Optimization. Use grid search or numerical optimization over (θ, ϵ) . Typical ranges: $\theta \in [0.001, 0.5]$, $\epsilon \in [-7, -1.5]$.

Q.4 Implementation Details

Focal ending. Set $\Delta = 0$ throughout (the focal point is like ignoring the end digits. In practice the perceived parameters are largely insensitive to this choice).

Polynomial degree. A cubic or quartic polynomial typically suffices for $\hat{F}_p$.

Price range. Focus on prices in a range where the density is reasonably high (e.g., \$2-\$15 for supermarket products). Exclude outliers.

Pre- vs. post-reform. For pre-reform estimation, just-below prices end in .99 and missing prices are .00 to approximately .28 (depending on θ). For post-reform, just-below prices end in .90 and missing prices are .00 to approximately .09.

Elasticity. In some specifications, ϵ is fixed at an externally estimated value (e.g., from demand estimation) and only θ is estimated. In others, both are estimated jointly.

Q.5 Interpretation

The estimated $\hat{\theta}$ represents the level of left-digit bias that *firms behave as if they believe consumers have*. This “perceived” bias can differ from the true bias estimated from demand data:

- If $\hat{\theta} < \theta^{\text{true}}$, firms underestimate left-digit bias and price as if the demand discontinuity is smaller than it actually is.
- If $\hat{\theta} > \theta^{\text{true}}$, firms overestimate the bias.

In this paper, I find $\hat{\theta} \approx 0.02\text{--}0.04$ from pre-reform pricing, while demand-side estimates suggest $\theta^{\text{true}} \approx 0.25\text{--}0.38$. Firms substantially underestimate left-digit bias, but their perceived bias is still positive. If we wish to keep the model with incorrect beliefs as a good positive description of how they price, we would interpret their actions as if they recognize that just-below pricing works, but not how well. However, as Section 5 concludes, such a model would be rejected in favor of heuristic-based pricing.