Learning to Set Prices in the Washington State Liquor Market

Yufeng Huang, Paul B. Ellickson and Mitchell J. Lovett*

October 8, 2018

[Preliminary and comments welcome!]

Abstract

How do new entrants learn about demand and adapt to new market conditions? We study retail prices in the Washington State liquor market where privatization led existing grocery chains to enter this market for the first time. We document large price changes across a broad range of products and provide novel evidence showing that these changes result from retailers learning about demand: prices absorb realized demand shocks, adjusting to better reflect demand primitives. We then estimate a structural model that imposes minimal assumptions on the optimality of observed prices. Comparing against the full-information optimal prices implied by the model, we find that learning continues to occur even years after entering the new market, and the limited demand information initially causes up to 13% lower profit compared to full information. Further, empirical pricing patterns suggest that retailers learn about both levels and the slope of demand, and learn from both own experience and practices of others.

*University of Rochester Simon Business School. We gratefully acknowledge numerous comments and suggestions from Kristina Brecko, Christopher Conlon, Ulrich Doraszelski, Ronald Goettler, Brett Gordon, Avery Haviv, Jihye Jeon, Przemyslaw Jeziorski, Sarah Moshary, Simha Mummalaneni, Harikesh Nair, Xiliang Lin, Marc Rysman, Katja Seim, Greg Shaffer, Thomas Wollman, Chenyu Yang, and Hongsong Zhang, as well as seminar and conference participants in the 2018 NYU IO Day, 2018 Summer Institute in Competitive Strategy, Northwestern University, 2017 SHUFE IO Conference, University of Chicago, and University of Rochester. Our empirical results are derived based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
1 Introduction

“Pricing is a big question mark, for everyone entering the spirits business in Washington... I sure don’t know what we’ll charge the consumer. There is going to be a lot of scrambling...”

– Alan Johnson, CEO of BevMo!\(^1\)

Canonical models in industrial organization, quantitative marketing, and macroeconomics characterize firms as setting prices with all available information. Empirically, however, managers’ responses to surveys and the prices they set are sometimes very different from what such models suggest. In surveys, experienced pricing managers often indicate that they do not (explicitly) know demand when setting prices (Noble and Gruca, 1999). In addition, as the macroeconomics literature documents, prices are often sluggish to respond to new information (Mankiw and Reis, 2002; Argente and Yeh, 2018). Further, recent empirical industrial organization literature also presents evidence that prices sometimes fail to account for differences across markets (DellaVigna and Gentzkow, 2017, among others) or to respond to large changes in market conditions (Arcidiacono et al., 2016). One possible conjecture to rationalize some of these pricing patterns is that it takes time for the firm to learn about new information. Nevertheless, there exists little direct evidence of firm learning and this conjecture is largely an open empirical question.

In this paper, we focus on a new market where entrants initially face demand uncertainty and study whether they learn about demand and set more-informed prices, and if so, how such learning takes place. In answering this question, we present new evidence that prices (1) respond to realized demand shocks in ways that diminish over time, and (2) become increasingly reflective of underlying demand fundamentals. We demonstrate using a structural model that, at the start of the market, demand information would have resulted in up to 13% improvement in firm profit. Finally, we show significant roles of prior experience and learning from others in rationalizing firms’ initial pricing mistakes, and discuss plausible mechanisms of how firms learn.

Empirically identifying post-entry learning behavior poses a steep challenge: most settings where such learning is likely to occur also present many confounds. Learning after disruptive innovations involves a complex innovation game between competitors (and potentially partners); learning when entering a new market involves gradual changes in customer behavior or market structure, new management routines for the firm, or complicated strategic choices beyond just pricing. Typically, the presence of many moving parts makes it difficult, if not impossible, to isolate learning about demand from a myriad of other, possibly more important factors. We identify a context for our study that provides an almost ideal, lab-like setting to examine how new entrants learn about demand.

We study pricing decisions made by established grocery retailers who enter the liquor market in Washington State after the privatization of the market. In Washington, existing products were previously sold by a state-owned chain, who committed to following a fixed-percent-markup pricing rule and did not set prices based on demand conditions. The well-anticipated privatization took effect in June 2012 and allowed larger retailers to obtain licenses to sell liquor. We argue and demonstrate that this setting has several features that simplify our inference problem markedly: customers with stable preferences for liquor, retailers with established positions and stable customer bases for which liquor does not appear to drive store visits, and a managerial decision, pricing, that has established routines that can directly extend to this new category. These aspects of the setting allow our inquiry to focus on the most critical uncertainty that retailers faced in setting prices – consumer demand for products. Our study focuses on whether, after entering the liquor market, these chains learn about demand over time and improve their pricing decisions.

We document large and heterogeneous price movements in the first two years after the privatization of liquor sales. Whereas median price across products drops by about 10% in the first year and remains largely stable afterward, prices for individual products change in heterogeneous directions and in large magnitude: most of the products change prices by at least 5% (up or down), and 20% of products change their prices by at least 15%. We present two pieces of descriptive evidence that suggest these price movements are due to retailers’ learning about demand. First, we
show that retail prices for a product respond to lagged demand shocks for the same product and that the rate of this response declines over time. This pattern is consistent with retailers learning from the realization of sales – i.e., they respond to sales shocks as new information, but the influence of this new information decreases as the retailer gains more information about preferences. Second, novel to the literature, we show that, across products, the correlation between observed prices and average quantities increases over time. This finding suggests that retailers learn about the demand for different products – i.e., they are able to identify products that consumers are willing to pay more (less) for and set higher (lower) prices correspondingly. We also provide evidence that these price movements are not reflective of several potential alternative explanations, including changes in retail outlets, assortment, competitive environment, consumer tastes or price search.

We then investigate more formally the degree to which firms are able to improve their performance by setting more-informed prices. To do so, we estimate a structural model of demand and cost primitives, imposing minimal assumptions on the optimality of the retailers’ pricing decisions, and compare its implied, optimal prices against the observed ones. We estimate a random coefficient demand model, incorporating standard aggregate- and micro-level moments. The resulting parameter estimates and output accords well with existing evidence from related liquor markets (Conlon and Rao, 2015; Miravete et al., 2017). Next, we estimate firms’ marginal costs based on the assumption that firms have reached full information in the last year of the sample. We then verify that these cost estimates are very close to the observed costs in Washington State before the market is privatized. After estimating demand and costs, we use the model to simulate prices set by full-information firms, and use these prices as a “normative” benchmark against observed prices.

We present three conclusions. First, retailers do learn about demand over time and prices become more similar to the optimal full-information levels. Second, early in the market, prices are set to sub-optimal levels compared to the (theoretical) perfect-information case. In particular, the lack of optimality corresponds to as much as a 13% loss in gross profit for retailers who are first-time liquor sellers. Third, the contrast between observed and optimal prices suggest that learning follows intuitive patterns. We find that the initial prices set by the retailers suggest they lack initial
knowledge about 1) demographics of liquor customers, 2) price elasticity of demand under the (relatively high) Washington tax regime, and 3) individual product demand in Washington. We also show that retailers not only learn from own experience but also from the behavior of retailers in other states.

Our main contributions are two-fold. First, we provide new evidence on firm learning and the impact of related practices on profits. We show that retailer strategies respond to past demand shocks in a manner consistent with learning about demand for specific products. We further establish that retailers set prices that increasingly capture demand, thereby suggesting that they obtain increasingly precise information about their products. To our knowledge, the second piece of evidence is novel to the literature. The closest related paper is Jeon (2017), who studies firm investment decisions with demand uncertainty in the container shipping industry. She measures investment costs directly, estimates a dynamic investment game with structural breaks in demand and firm learning, and finds that firms learn more from recent realizations of demand shocks than from those in the distant past. Our paper is different in that we provide direct, descriptive evidence that firm learn about demand and descriptively characterize in what ways learning occurs. Also different from Jeon (2017), we focus on learning about demand in a new (and stable) market while she focuses on firm adapting to changes.

In addition, our paper is also closely related to Doraszelski et al. (2018), who show that after the market for frequency response opens (in the UK electricity system), prices appear random at first and evolve to patterns that can be rationalized as stationary equilibrium. They focus on adaptation of equilibrium play whereas our paper focuses on firm learning about demand. Finally, our paper is also related to Hitsch (2006) and Covert (2015). Hitsch (2006) documents that, in the ready-to-eat cereal market, many products remain in the market for a long time despite making low sales, and estimates a structural model in which forward-looking manufacturers learn about the demand for their products. Covert (2015) investigates firms’ learning about shale productivity in the Hydraulic fracturing industry. He shows that firms respond to public reports about production output and adjust their production strategies, and finds that firms overweight own signals and incorporate
public signals sub-optimally.

We also contribute to the discussion regarding heterogeneity in management practices within firms. There are broadly two approaches in this literature. In one approach, Noble and Gruca (1999) and Bloom et al. (2017) base their evidence on surveys and show that there is heterogeneity in firms’ reported business practices. Bloom et al. (2017), among others, show that such differences explain a sizable fraction of the productivity differences across firms. In the other approach, Goldfarb and Xiao (2011) and Hortacsu et al. (2017) show that there is heterogeneity in observed strategies. Whereas the strategies for some firms better fit an economic model (e.g. in Hortascu et al., larger firms more closely follow model predictions), the behavior of other firms is more difficult to rationalize. Our approach lies closer to the latter, and we add to this literature by showing that pricing strategies can be sub-optimal when a firm is inexperienced – and more importantly, this lack of experience is not only an outcome of lack of information but rather lack of understanding of existing information.

Secondarily, we also add evidence to recent descriptive studies of uniform retail prices. DellaVigna and Gentzkow (2017) and Hitsch et al. (2017) document that retail prices are geographically uniform and discuss the extent to which they are set sub-optimally by retailers who ignore heterogeneity across markets.\(^2\) We confirm the same case in Washington State liquor market. Also, we show large gains from obtaining information about demand, which speaks to the hypothesis in DellaVigna and Gentzkow (2017), who propose that one way to rationalize the uniformity in prices is that locally-optimal prices demand more information about local demand and is costly to maintain for the firm.

Finally, our paper also contributes to the research and evidence in recent studies about the retail liquor industry. Seo (2016) studies consumer store choice before and after the privatization of Washington State liquor market and quantifies the welfare impact of one-stop shopping. Illanes and Moshary (2017) leverage Washington State’s 10,000 square-foot minimum required

\(^2\)Li et al. (2017) and Adams and Williams (2017) find prices in zones and rationalize this pricing pattern as a way to soften local market competition. Chintagunta et al. (2003) also documents uniformity in prices and quantifies the gains from pricing differently across markets.
retail space for liquor vendors, as a regression-discontinuity design, and study the effect of entry on prices and product assortments. Conlon and Rao (2015), Aguirregabiria et al. (2016), and Miravete et al. (2017) study different aspects of liquor regulation policy. Conlon and Rao (2015) study the “post-and-hold” policy as a collusive instrument for the wholesalers and investigate the welfare improvements (and redistribution) of an alternative tax policy. Aguirregabiria et al. (2016) investigate counterfactual regulation, tax, or competition regimes in the Ontario wine market, highlighting the importance of spatial differentiation. Miravete et al. (2017) study the welfare impact of state-imposed constant retail markup and find that the single-markup policy decreases but also re-distributes consumer welfare.

Section 2 briefly describes the privatization of the Washington liquor market. Section 3 presents the data, and Section 4 provides key descriptive evidence about retailer learning. Section 5 then estimates a structural model and backs out demand and costs. After the structural estimation, Section 6 describes firms’ learning process from the structural estimates. Section 7 concludes the paper.

2 Privatization of liquor sales in Washington State

Before June 2012, a state-owned chain operated by the Washington State Liquor Control Board (WSLCB), "Liquor and Wine," monopolized the off-premise liquor market in the state. The state-owned chain committed to set prices at a 51.9% markup over the wholesale price. In Fiscal 2011, WSLCB directly operated 166 stores in cities and contracted with private owners who operated 162 stores in rural areas. Total sales revenue for alcoholic beverages amounted to $888 million in 2011.

On November 8, 2011, Initiative 1183, which mandated the privatization of the Washington State liquor business, was passed with 59% of voter in favor. I-1183 mandated: (1) that WSLCB

---

3 This percentage markup was previously 39.2% and increased to 51.9% on August 1, 2009. Source: WSLCB press release, May 6, 2009.

must exit the liquor business before June 1, and auction off its inventories, (2) that private retailers are eligible to carry and sell liquor products upon obtaining a license, starting from June 1, 2012, (3) that private retailers must pay 17% of liquor revenue to the state as part of the licensing fee, and (4) that retail off-premise sales taxes will increase from 10% to 20.5% (excise tax, at $3.77 per liter, stays unchanged). In particular, the Initiative requires that retail licenses can only be issued if the store has at least 10,000 square feet of floor size. Before I-1183, two other initiatives to privatize the retail liquor market were voted down in 2010.

After the privatization of liquor sales, the WSLCB began issuing retail licenses on March 1, 2012. We observe that grocery stores and other retail chains (who do not specialize in liquor sales) immediately became dominant players starting from June 1. That grocery retailers dominated the market is consistent with Article (3)(b) in I-1183, which prioritizes the issuance of retail liquor licenses for “existing grocery premises licensed to sell beer and/or wine.” Also, the press reported that many of the state stores remained in business but are operated by independent owners.

A report by the State shows that post-privatization, total sales volume increased by about 20%, average prices increased by about 8%, and the total number of liquor-selling stores increased by 3.27 times (Washington-State, 2015). This finding is consistent with the results in Seo (2016) and Illanes and Moshary (2017).

3 Data and summary statistics

3.1 Main data and sub-sample

Our primary data source is the Nielsen Retail Measurement Services (RMS) Dataset.\footnote{We thank the Kilts Center for access to the data. Nielsen retains copyright of the data: Copyright © 2018 The Nielsen Company (US), LLC. All Rights Reserved.} For participating chains, the Nielsen RMS dataset contains information about price, quantity sold, and feature and display, recorded at the store-UPC-week level, for a broad set of consumer packed goods. For our study, we focus on the liquor category in Washington State, from June 2012 (the
month of privatization) to December 2016. To limit the complexity of our analysis (especially in
the structural model), we further focus on the broad whiskey category, which consists of whiskey,
bourbon, scotch and rye, a set of reasonably closely substitutable products. This sampling proce-
dure yields a dataset of 6,288,941 observations at the UPC-retailer-store-week level. In particular,
the sample contains 724 unique UPCs (UPCs as product name - size combinations; there are 635
unique product names), 625 stores from six retail chains, and 240 weeks between June 2, 2012,
and December 24, 2016.

We further restrict our attention to a smaller subsample through the following two cuts. First,
we focus on stores that sell a positive quantity in at least 95% of all weeks. This step selects 561
out of 625 stores and removes 5.9% observations from the overall sample. We note that most of the
cases in which quantities go to zero are not due to the decision to carry or drop liquor, but instead
to entry or exit of the store as a whole. Therefore, we condition on the set of stable stores.

Second, we focus on “core assortments” defined as those, within a given retailer, that first
appear before December 2012, last appear after March 2016, and is present for at least 25 weeks.
This sample-selection rule means that we condition on a set of products that the retailers carry
from essentially the beginning to the end of the sample, abstracting away from new products,
discontinued products, or products that are only occasionally available. This step selects 276 out
of 724 UPCs in the full sample. Although it may appear that many of the UPCs are eliminated in
this step, the products we drop out in this step only account for 15.8% of the sample size of stable
stores and 11.4% of the total revenue from these stable stores. We discuss the choice of focusing on
these core products in more detail in Section 3.4.2. After these sample selection steps, our sample
contains 4,985,621 observations, from 276 products, six retailers, and 561 stores.

---

6 Only for Washington State.
7 The sample we discuss has already excluded two grocery retailers that entered in small scale in some parts the
market in 2016 (24,225 observations, 0.3% of the sample size), and two mass merchandisers that operate in small
scale in the liquor market. One of the two mass merchandisers only have 650 observations (0.01% of the sample size),
and the other has 103,942 observations (1.6% of the sample size). These observations are either anomalous or involve
different retail formats that are small in the liquor market.
8 By the account of Section 2, our sample contains 55% of all the liquor-selling stores in the state.
3.2 Auxiliary data

We supplement the main data with a few auxiliary datasets. First, in some of our descriptive analyses, we compare Washington to other states that have allowed retail grocery liquor sales for as long as several decades. We identify fifteen such states in the Nielsen RMS data and use these as a kind of “placebo test.”\textsuperscript{9} In this auxiliary sample, retailers have generally been in the liquor business for a long time in comparatively stable market conditions, so that one does not expect them to exhibit strong learning about demand. Therefore, we contrast our evidence about retailer learning in Washington with evidence from those states.

Next, we use public price data from Washington and Oregon to help estimate or validate retailer wholesale prices in our structural analysis. Firms in both Washington (before privatization) and Oregon (throughout the sample period) practiced fixed markup policy.\textsuperscript{10} Thus, one can directly infer wholesale prices from observed retail prices in both states. On the one hand, we use average wholesale prices over time in Oregon to supplement our model and data in Washington’s post-privatization market to infer wholesale prices. On the other hand, while we do not impose wholesale prices stay the same before and after the privatization in Washington, we compare our implied wholesale prices with the pre-privatization data in Washington and show that they are similar.

Finally, we also use the Nielsen Consumer Panel Data between December 2009 and December 2016 to examine consumer behavior in greater detail and to construct micro-moments (Petrin, 2002) for use in structural demand estimation. Within the panel data, 2,952 households reside in Washington State, 1,177 of them ever purchased liquor in the sample period, and 498 households purchased at least five times (in all stores). Hence, because the household panel data contain very infrequent liquor purchases, we rely on the RMS data for our primary analyses.

\textsuperscript{9}These states are, in descending order of total liquor sales volume within our data: California, Arizona, Louisiana, Texas, New Mexico, Nevada, Nebraska, Wyoming, South Dakota, Colorado, Arkansas, Delaware, Maryland, North Dakota, and Washington, D.C.

\textsuperscript{10}For the state of Oregon, ORS 471.745 mandates a fixed markup. Various sources indicate that the markup is set at 106%. See Governor’s Task Force, 2003, “Final Report on the Oregon Alcohol Beverage Industry” (page 25). Also see The Romain Group, LLC, 2014, “Written Comments on Draft Ballot Title for Initiative Petition No. 57” (page 2).
3.3 Price changes after the privatization of liquor sales

3.3.1 At which geographic level are prices set?

As our focus is on learning about demand, it is important to establish where that learning might manifest itself and the degree of price variation observed over time. Consistent with DellaVigna and Gentzkow (2017) and Hitsch et al. (2017) we find that, for a given product at a given point in time, there is little price variation within a chain across stores and markets. We document the details in Appendix A. However, there are sizable price differences across retailers and over time. Therefore, we treat pricing decisions as if they are made on the retail chain level and investigate changes in these prices over time (and the extent to which such changes are driven by learning about demand).

3.3.2 How much do prices change over time?

We first examine how the overall Washington liquor price level changes over time. We focus on the 750ml bottle size and match 65 products that are available both before and after the privatization. We compare volume-weighted average prices before and after, and find that, at the privatization, the average price increases by 41%, from $14.80 to $20.83. Prices then drop to about $20 in the first year and stay stable.

Second, we examine the magnitude of price changes relative to the prices at the privatization. For each product-retailer, we first calculate the percentage changes in its prices (averaged across markets) relative to the initial price, \( \left( \bar{p}_{jrt} - \bar{p}_{jrt0} \right) / \bar{p}_{jrt0} \). We then plot the median, 25th/75th percentile, and 10th/90th percentiles of the normalized prices over time. Figure 1 shows the changes of these quantiles: In Washington State, prices undergo dramatic changes, and the paths are different across products. For example, in mid-2014, 25% of products are priced at least 8.6% below their starting price (at the privatization), and 25% of products are priced at least 6.3% above their starting price. In addition, much of the price movements happen in the first two years, with the distribution of price changes becoming stable afterward. Although this finding is not direct evidence

---

11We use each product’s total sales quantity in the second half of 2016 as weight, and fix these weights over time.
of learning per se, such dramatic price changes are consistent with retailers learning about demand and adjusting (the distribution of) prices. Appendix Figure 16 presents another way to visualize the price paths, plotting the relative prices (defined above) over time for all products separately.

3.3.3 Changes in promotion policy

It is also important to consider whether promotion policies evolved over the sample period, perhaps contributing to or explaining the price changes. To this end, we first define regular price as the 90th percentile of prices for a given retailer-market-product in a given quarter, and define promotion instance as when the distance between list price and regular price is beyond 5% below the regular price (and promotion depth accordingly). Promotion as defined occurs at a frequency of 22.2%. When a product is on promotion, the average promotion depth is 12.8% of the regular price (25/75th quantiles at 7.2% and 17.1%).
Figure 2: Left: shelf price and regular price (dash), Right: promotion frequency and depth (dash)

Notes: Left panel: solid line is the average shelf price; dashed line is the average regular price defined as the 90% quantile of shelf price in each quarter. Right: solid line is promotion frequency defined as the share of products having at least 5% differences between regular and shelf prices; dashed line as promotion depth defined as the percentage differences between shelf price and regular prices conditional on product being on promotion.

Figure 2 presents time trends in shelf price, regular price, promotion frequency and promotion depth. We find that the main variations in promotion frequency seem to be seasonal variations, and promotion depth increases from about 12% to 14% over the sample period. As a result, the gap between regular and shelf price roughly stays stable in the sample period, whereas the main variations of interest seem to be the variations in regular prices over time. For simplicity, and also because the shelf price is clearly defined by the data while the regular price relies on additional definitions, we model the shelf price.

3.4 Do price changes reflect phenomena other than learning about demand?

In this section, we examine a few factors (other than prices) that might change because of the privatization of liquor sales. Our aim is to rule these out as alternative dimensions on which firms might be updating their beliefs.
3.4.1 Changes in the set of stores that carry liquor products

One might expect retailers to make decisions about which stores to carry liquor and for such decisions to change as they learn about demand. We examine whether the set of stores selling liquor changes over time and rule out this conjecture. We compare the number of distinct stores selling grocery items for a given chain, against the number of stores selling liquor for the same chain. Appendix Figure 14 report these numbers over time and by chain. For the set of chains in the liquor market, almost all stores carry liquor as long as they remain in business. We conclude that which stores carry liquor is not a relevant decision for the chain or store managers.12

3.4.2 Changes in liquor product assortments

Alternatively, one might expect retailers to make decisions about which assortments to carry and that such decisions could also be an outcome of learning. We examine whether the decision of which products to carry changes over time and gauge the overall magnitude of assortment changes. As mentioned in Section 3.1, among the 724 products (as product name - size combinations, sizes can be 375ml, 750ml or 1750ml), only 276 products are “core” in the sense that they are carried by the retailer between privatization and the end of 2016. The “non-core” products represent 15.8% of the sample and account for 11.4% of the overall revenue. Appendix Figure 15 shows these core products account for most of the revenue in all time periods.

Although firms may adjust assortments over time, we conclude that these decisions are only relevant for low-revenue products. Thus, we focus our inquiry around the pricing of core products.13

12One might be concerned that stores that sell little quantity might be mis-recorded as not carrying liquor, as liquor is a slow-moving category. We collect data on the identity of off-premise liquor license holders by year from the Washington State Liquor Control Board, cross-check the number of distinct license holders by retailer by year, and find that the number of license exactly match the number of stores in the Nielsen data. This additional check indicates that there is no measurement error in the store identity.

Figure 3: whiskey sales revenue decomposition, across retailers

Notes: Decomposition of liquor sales revenue across six focal retailers in Washington State, focusing on the set of core products.

3.4.3 Changes to the market structure

One might further expect the market structure to change over time as new entrants come to the market and compete against the initial set of retailers. Figure 3 shows that the total revenue and revenue composition among the six focal retailers are stable over time: it seems that the market structure is stable as each retailer occupies a stable share of the market.

Nielsen RMS data does not cover all retailers in the market. However, the Homescan data do cover, for a small set of consumers, expenditure in all retailers. We use these data to measure the revenue of the six retailers in our main data relative to the rest of the market. We find that after the market is privatized, the six retailers collectively take 37% of the total revenue on average. In addition, their revenue share in the market is stable over time, as shown in Appendix Figure 18. These descriptive evidence suggest that market structure stays stable with respect to retailers outside of our sample.

Finally, although the retail market structure remains stable in the sample period, one might expect the wholesale market to change over time. While we do not have data about wholesale
contracts, industry publications reveal that the liquor wholesale market was dominated by two national wholesalers and the market structure did not change with privatization.\footnote{These two wholesalers are Southern Glazer’s Wine and Spirits, and Young’s Market – Columbia Distributing. See Erickson, “A Status Report on the Implementation of Measure 1183”. Extracted from http://www.healthyalcoholmarket.com/pdf/Alcohol_Deregulation_by_Ballot_Measure_in_Washington_State.pdf in July 2018.} We examine historical data on wholesale license ownership (from WSLCB) and find that, in the period between 2014 and 2016, license distribution is stable among distributors in Washington State,\footnote{In the three-year period, Southern owns between 12 and 15 wholesale licenses, Young-Columbia owns between 31 and 33, and the next largest distributor, The Odom Corporation, owns between 3 and 5 licenses.} supporting the view that the wholesale market is stable over the sample period.

### 3.4.4 Changes in shopper behavior

Furthermore, one might expect consumer behavior to change with privatization. In particular, if consumers gradually learn about the privatization event, or about which retailers carry liquor post-privatization, or develop their preferences over time as new product varieties emerge in the market, a rational and fully-informed retailer might adjust prices over time as demand evolves. This is to say, consumer learning might be able to rationalize price adjustments after liquor privatization even when firms are fully informed.

However, we show in Figure 3 that sales are stable over time in the sample period, consistent with stable demand. We further show, in Appendix Figure 17, that several measures of consumer behavior using the Nielsen Homescan data are stable in the early periods after the privatization.\footnote{We also note that expenditure and other household choices, measured by the Homescan data, fall slightly in the last 1-2 years. In contrast, the point-of-sales data (RMS) do not show such a pattern. We examine the changes in household demographics and find a slight decrease of income among active shoppers. Thus, we do not interpret the slight decrease in household behavior at the end of the sample as representative of the market.} Our structural estimates presented in Section 5 directly show that demand is stable over time.

### 3.4.5 Competition between chains

Finally, one might expect chains to compete with one another for consumer traffic in the Washington liquor market. We argue that competition is not first-order in the context based on three
observations.

First, one might imagine that firms will use the liquor category as a “traffic driver”, i.e. set low liquor prices to drive grocery traffic. We examine consumers’ choices of primary store and find that only about 4% consumers switch primary stores in each year, and importantly, the share of store-switchers is stable at and after liquor privatization. In addition, we show that liquor prices are too high at the start, suggesting that firms did not try to underprice liquor to drive store traffic.

Second, there are few dedicated liquor trips and almost all liquor-shopping trips involve significant grocery expenditures. We document that the median share of liquor expenditure per trip is 0.1% among all trips, or 28.7% among trips with liquor purchases. This finding suggests that liquor shoppers are part of the regular grocery-shopper population, instead of being a new and isolated consumer segment over which the firms might compete.

Third, we show that there are limited overlap for grocery shoppers among the set of focal retailers. Among household-year, the median share of expenditure at the primary chain (i.e. the chain where the household spends the most in that year) is 80%. In contrast, the median share of expenditure at the secondary chain (i.e. chain with the second-highest expenditure) is 17%. Given the limited shopper overlap, and that few grocery shoppers buy liquor, it seems unlikely that chains compete in liquor prices. In addition, we provide alternative evidence in Appendix B, showing that the sales of a product does not decrease if other retailers in the same market carry the product or promote the product. These evidence further suggest that substitution between chains for a liquor product is not detectable in our sample. Therefore, it is a reasonable approximation that grocery retailers set liquor prices for existing customers as if they are a monopolist on this customer base.

4 Learning about demand: empirical evidence

In this section, we provide two key pieces of evidence, showing that retailers in Washington State learn about demand for liquor and adjust their prices. First, we show that prices adjust according

\[\text{See Appendix Figure 17}\]
to past demand shocks when retailers start selling liquor in Washington, and that, over time, they adjust less and less to such shocks. This is consistent with a rational learning model where a retailer who does not know about demand will adjust prices according to temporary demand shocks, whereas the retailer who later knows demand does not adjust prices to such shocks. Second, we show that the price of a product increasingly reflects its demand level (i.e., the intercept in the demand function). This implies that retailers gain knowledge about which products sell better in Washington and are able to adjust their prices to reflect this information.

Although our main focus is to test for learning in Washington, we present the same set of evidence in other states as a “placebo” test. In other states, retailers have been selling liquor products since long before June 2012. Therefore, we expect that our evidence for learning should not appear in these placebo tests.

4.1 Do prices adjust based on past demand shocks?

If a retailer learns about demand by observing (noisy) realizations of demand shocks, she will initially adjust prices according to those shocks, but later (as she has already learned) cease doing so. We test whether retailers adjust prices according to demand shocks when they enter the market and whether they do so to a lesser extent as they gain experience.

We first aggregate our data to the product-retailer-month level to reduce the effect of temporary price promotions and to work at the geographical aggregation where prices are set. Denote $j$ as a product, $r$ as a retailer and $t$ as a month, we estimate a linear model of current price on 1-month lagged quantity, controlling for current quantity and lagged prices:

$$
\log (p_{jrt}) = \beta_t \log (q_{jrt-1}) + \\
\rho \log (p_{jrt-1}) + \alpha^{-1} \log (q_{jrt}) + \psi_{jr} + \phi_t + \eta_{jrt}.
$$

(1)

Our key parameter of interest is the sensitivity of the current price to $q_{jrt-1}$, which is the units sold for product $j$ by retailer $r$ in the previous month, $t - 1$. We allow $\beta_t$ to take a different value for
each half-year. Learning would predict that these sensitivities should be initially positive, but then decline in magnitude over time. We include product-retailer fixed effects, \( \psi_{jr} \), that capture stable product-retailer characteristics that could affect price levels, time fixed effects, \( \phi_t \), that capture common variations price across time, past logged prices (with parameter, \( \rho \)) which captures serial correlation in price beyond these fixed effects, and logged current quantity (with parameter, \( \alpha^{-1} \)) that captures the correlation between current quantity and price,\(^{18}\) and can be roughly thought of as the inverse price elasticity.

For estimation, we take first difference to net out the fixed effects. After first differences, \( \Delta \log (p_{jrt-1}) \) and \( \Delta \eta_{jrt} \) are mechanically correlated because of the first difference, and we correct for such correlation using Arellano and Bond (1991) instruments. To instrument the first lag of price, we use the third lagged price difference, \( \Delta \log (p_{jrt-3}) \), which guards against potential additional serial correlation in \( \eta_{jrt} \) that might infect the second lag. Such serial correlation could potentially arise because retailers are unwilling to change price right after a price change due to reasons such as menu costs.

Table 1 presents these estimation results and Figure 4 plots the coefficients. We estimate and plot separately Washington and the placebo states where retailers are well-experienced in selling liquor. We check the first stage and find that the Arellano-Bond instruments are strong and the coefficients on control variables have plausible signs and magnitudes. As main results, we find that in Washington, prices respond to the previous month sales quantity positively, but the effect decreases over time. Just after the privatization of liquor sales, an increase from the mean of 10% in past sales increases prices in the next month by 0.16%. Four years later the responsiveness decreases by half to only 0.08%. Hence, retail prices incorporate past sales shocks but decreasingly so, consistent with retailers learning about demand by incorporating information contained in the realized sales quantities. This pattern for Washington State stands in sharp contrast to what we find for the states where retailers already had extensive experience in the liquor market. In these states, the responsiveness to sales shocks is much lower and all coefficients but one are insignificant.

\(^{18}\)Current quantity is correlated with past quantity and thus cannot be omitted from the regression.
Hence, with greater experience, prices do not respond to past sales shocks.\(^{19}\)

Figure 4: The response of current price to lagged quantity, Washington (left) and other states (right)

Notes: Estimates \(\beta_\tau\) from Table 1. See table notes for more detail. Confidence intervals are two standard error.

4.2 Do prices increasingly reflect demand?

One of the early motivations for price endogeneity bias is that prices are set in part to reflect the underlying quality of the product (Trajtenberg, 1989; Berry, 1994). As a result, if cross-sectional variation were used to estimate price sensitivity, the price sensitivity would be biased toward zero. We extend this logic to the setting where the price-setter is learning about demand.

We argue that if retailers are uncertain about demand initially but learn about it over time, prices should evolve to increasingly reflect demand as the learning process unfolds. Prices for products that have surprisingly high-demand (relative to the prior belief) will rise to reflect that high demand, and prices for surprisingly low-demand products will fall. The end result is that prices and the demand levels of products should become more highly correlated as the retailer

\(^{19}\)Except for the second half of 2012.
Table 1: The response to current price to lagged quantity, Washington and other states

<table>
<thead>
<tr>
<th></th>
<th>WA: D.logprice first stage</th>
<th>Ctrl: D.logprice first stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>past month D.logprice</td>
<td>0.198*** (0.055)</td>
<td>0.182*** (0.037)</td>
</tr>
<tr>
<td>L3D.logprice</td>
<td>-0.109*** (0.005)</td>
<td>-0.067*** (0.002)</td>
</tr>
<tr>
<td>D.log quantity</td>
<td>-0.038*** (0.001)</td>
<td>-0.003*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2012.5)</td>
<td>0.016*** (0.003)</td>
<td>0.005*** (0.002)</td>
</tr>
<tr>
<td>D.past month log quantity (2013)</td>
<td>0.013*** (0.003)</td>
<td>0.003* (0.002)</td>
</tr>
<tr>
<td>D.past month log quantity (2013.5)</td>
<td>0.014*** (0.003)</td>
<td>0.003*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2014)</td>
<td>0.007** (0.003)</td>
<td>0.001*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2014.5)</td>
<td>0.008*** (0.002)</td>
<td>0.002*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2015)</td>
<td>-0.001 (0.002)</td>
<td>0.002*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2015.5)</td>
<td>0.005** (0.002)</td>
<td>0.002*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2016)</td>
<td>0.004* (0.002)</td>
<td>0.001*** (0.001)</td>
</tr>
<tr>
<td>D.past month log quantity (2016.5)</td>
<td>0.008*** (0.001)</td>
<td>0.001*** (0.001)</td>
</tr>
<tr>
<td>constant</td>
<td>-0.039*** (0.006)</td>
<td>0.020*** (0.003)</td>
</tr>
<tr>
<td>yearmonth FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Rsq.</td>
<td>0.002 0.122</td>
<td>0.054</td>
</tr>
<tr>
<td>obs.</td>
<td>3.3e+04 3.3e+04 2.2e+05 2.2e+05</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Column 1 reports estimates of Equation (1), estimated using Washington data (first two columns) and compared against data from other states (last two columns). Lagged price difference is instrumented by the third lagged price difference, and Column 2 reports first stage results (dependent variable is lagged log price difference). Columns 3 and 4 estimates the same specification on other states. All standard errors are robust and clustered on the product-retailer level. *, **, and *** indicate significance at the 90, 95, and 99 percent. F-test for excluded variables in first stage: 392 for Washington, 912 for other states.
learns about demand. Hence, learning should be reflected in a pattern of increasing correlations between prices and the underlying demand for the product.

This intuitive argument can be illustrated with a simple model has linear demand for product $j$ in retailer $r$ and month $t$ determined by price and an intercept $\delta_{jr}$,

$$q_{jrt} = \delta_{jr} + \alpha p_{jrt}. \quad (2)$$

Assuming profit maximization and zero marginal costs, the optimal price would be set to $p_{jrt} = -\frac{\delta_{jr}}{2\alpha}$. For now, imagine the slope $\alpha$ is known to the retailer from serving these customers for many years in other categories, but the intercept $\delta_{jr}$, which is specific to the new market, is not fully known.\textsuperscript{20} As retailers learn about demand, their belief $\delta_{jr}$ is closer to the truth, and so are the prices to the optimal, full-information ones. Hence, as retailers learn over time, the correlation between the prices and this underlying demand primitive would on average increase.

This observation and illustration motivates a new descriptive test of firm learning about demand. The concept is to estimate the cross-sectional relationship between prices and the sales for a product in each period and evaluate how this correlation changes over time. If learning occurs, we would expect the correlation to increase over time and otherwise be (potentially noisy, but) stable. To implement this test, we focus on quantity and price data on the product-retailer-week level. Pooling these data to half-year time windows (denoted $\tau$), we estimate the linear regression between quantity and price by ordinary least squares (OLS)

$$q_{jrt} = \bar{\delta}_{rt} + \alpha_{\tau} p_{jrt} + \eta_{jrt}. \quad (3)$$

With the retailer-week fixed effects $\bar{\delta}_{rt}$ in Equation (3), the remaining variations left in the error term is only across-product, i.e. $\eta_{jrt} = \delta_{jr} + \xi_{jrt} - \bar{\delta}_{rt}$ (where $\xi_{jrt}$ represents idiosyncratic demand shocks). Thus, besides capturing the slope of demand, the regression coefficient $\alpha_{\tau}$ might also

\textsuperscript{20} Although this illustration assumes knowledge of price sensitivities and zero marginal costs, these are not necessary for the main conclusion to hold.

22
reflect the retailer’s belief about the product quality, through the covariance between \( \delta_{jr} \) that is left in the error term, and some belief about it that are reflected by the observed prices. We report the estimated \( \alpha_{\tau} \)'s and their standard error for the Washington sample in Figure 5. We also examine retailers in other states as a placebo test: in that case, controlling for retailer-state-week fixed effects in a similar regression as (3).

In Washington, we find that retail prices are increasingly correlated with demand intercepts, reflected by the way \( \hat{\alpha}_t \) increases between 2012 and 2015. The relationship then stabilizes for the last two years. If the underlying price sensitivities are stable over time this finding suggests that retailers set prices with more and more information about the demand, so that the price sensitivities are biased more and more toward zero. In contrast to this pattern, we find that the other states exhibit no meaningful increases in \( \hat{\alpha}_t \) and that it is instead stable over time. This finding is consistent with these experienced retailers exhibiting no meaningful learning about liquor demand. Both of the plots support the conjecture that retailers in Washington learned about demand after deregulation.

We note that we return to this concept for testing learning in section C. In that section, we estimate our structural demand model parameters, which includes parameters for \( \delta_{jr} \). We then correlate those parameter estimates with the observed prices to see whether this correlation increases over time. Our results are consistent with the descriptive analysis presented here.

### 5 Model: demand and costs

We characterize the primitives of demand and costs with the goal of quantifying the effect of limited information on retailer profits. Demand for liquor is characterized by the random coefficient logit model, and estimated with standard nested-fixed point methods that incorporate micro-moments (Berry et al., 1995; Nevo, 2001; Petrin, 2002). We then estimate costs using the period between March 2016 and September 2016 – the last 6 months in the Washington sample.\(^{21}\) After reviewing the estimates, we present evidence that the model fits the data well and produces sensible

---

\(^{21}\)Note these are the last two quarters prior to the observed price decrease. As noted previously, we do not include the data during or after that unexplained price drop.
Figure 5: Coefficient estimates of $\alpha_{\tau}$ from Equation (3), Washington (left) and other states (right)

Notes: We report coefficient estimates of $\alpha_{\tau}$ from Equation (3) where $\tau$ represents half-year periods. The regressions control for retailer-state-week fixed effects but leave product fixed effects in the residual to explicitly capture the covariance between price and product fixed effects in the coefficient estimates. Confidence intervals are two standard error.

implied markups. We then characterize the variation in the implied marginal costs.

### 5.1 The demand for liquor

Consumer $i$ in market $m$ in month $t$ comes to retailer $r$ to buy groceries, and derives utility from purchasing liquor $j$:

$$u_{ijrmt} = \gamma_i + \alpha_i p_{jrm} + \delta_{jrm} + \lambda_{rt} + \xi_{jrm} + \epsilon_{ijrmt}. \quad (4)$$

In the above, $\gamma_i$ and $\alpha_i$ are household-specific intercepts and price coefficients, capturing heterogeneity in the tastes for liquor and sensitivities to prices across households. $x_{jrm}$ are observed promotion variables that vary over time.\(^{22}\) $\delta_{jrm}$ are product-retailer-market fixed effects capturing tastes to different products which could differ across shoppers going to different retailers or are

\(^{22}\) These are indicator variables of whether the product is on feature or display, or feature/display status unknown.
in different markets. $\lambda_{rt}$ are retailer-time fixed effects capturing changes over time in the demand for liquor in grocery stores, or changes in market positions for different retailers (in the liquor market). $\xi_{jrmt}$ are unobserved characteristics or demand shocks. $\epsilon_{ijrmt}$ are type-1 extreme value utility shocks. If the consumer does not buy any liquor in the given trip, her utility is normalized to $u_{i0rmt} = \epsilon_{i0rmt}$.

The consumer chooses among products in a given retailer, i.e. from choice set $J_{rmt}$. This choice set assumes that the liquor category does not drive store choice, and the retailers act as monopolists over their own store traffic. As noted earlier, this assumption is motivated by multiple corroborating observations including that consumers’ grocery shopping patterns after liquor deregulation are stable, and when a consumer makes a liquor purchase, she typically spends more on other grocery items than liquor on the trip, and the degree of overlap between retailer cross-shopping is limited. With these assumptions, the market share within a retailer-market is an integral of logit choice probability over random coefficients

$$s_{jrt} = \int s_{ijrmt} dF(\alpha_i) = \int \frac{\exp(\gamma_i + \alpha_i p_{jrt} + x_{jrm} \beta + \delta_{jrm} + \lambda_{rt} + \xi_{jrmt})}{1 + \sum_{j' \in J_{rmt}} \exp(\gamma_i + \alpha_i p_{j'rmt} + x_{j'rmt} \beta + \delta_{j'rmt} + \lambda_{rt} + \xi_{j'rmt})} dF(\alpha_i, \gamma_i).$$

(5)

To capture heterogeneous tastes, we follow Conlon and Rao (2015) and Miravete et al. (2017) and parameterize the household intercept as a function of log household income and an independent normal random draw $v_i$,

$$\gamma_i = \gamma_0 + \gamma_1 \log(y_i) + \sigma_v v_i,$$

(6)

and the price coefficient as a function of log household income

$$\alpha_i = \alpha_0 + \alpha_1 \log(y_i).$$

(7)
5.2 Identification

5.2.1 Price coefficient

We estimate model parameters by a set of moments enforcing that demand shocks $\xi_{jtmt}$ are conditional mean zero, given instruments $z_{jtmt}$ (including non-price covariates, fixed effects, and excluded instruments for price and random coefficients):

$$\mathbb{E} [\xi_{jtmt} | z_{jtmt}] = 0. \quad (8)$$

Despite the inclusion of fixed effects, retailers might set prices based on private information not directly controlled for by these covariates: for example, the retailer might set higher prices for certain products in markets or time periods with high demand, in which case prices will be correlated with $\xi_{jtmt}$ and the price coefficient estimate will be biased.

To address this concern, we construct price instruments similar to Conlon and Rao (2015) and Miravete et al. (2017). For each product, we take average prices across all states other than Washington and use them as instrument for retail prices in Washington. The prices in other states likely capture wholesale price variations that are common across states but do not correlate with demand shocks in Washington (after controlling for the above fixed effects). One example of wholesale price co-movement is that prices of Scotch whiskey move with the USD-GBP exchange rate, as illustrated in Figure 19 in the Appendix.

5.2.2 Random coefficients

We identify the random coefficients by combining the “BLP” instruments with additional micro moments (Petrin, 2002). First, we count the number of products available in each retailer-market-month, which are typically referred to as “BLP” instruments after Berry et al. (1995). Variations in the market shares of the focal product in response to changes in the number of products identify the substitutability to other products versus to the outside option, which is captured by the random intercept $\gamma_i$. 
Second, we construct two sets of micro moments, using the Nielsen Homescan panel data, to help identifying how the demand intercept and price coefficient vary with log income. Specifically, we divide annual household income (in thousand dollars) into three bins $I_b$: $(0, 42.5]$, $(42.5, 85]$, $(85, \infty)$. For Washington households in Homescan who visit the six focal retailers, these three income bins roughly divide them into three equal-sized groups.

Next, for each income bin, we compute the average probability of buying liquor given trips to the retailer,\(^{23}\) as well as the average price paid given liquor purchases. Then, for each set of parameters $\theta$ in the structural model, we compute the average probability of choosing the inside good,

$$\bar{s}_{b jrmt}^{b} = \frac{1}{N_b} \sum_{i \in I_b} \sum_{j \in J_{rmt}} s_{ijrmt}(\theta)$$

(9)

and the average price paid given purchase of liquor:

$$\bar{p}_{b jrmt}^{b} = \frac{1}{N_b} \sum_{i \in I_b} \left( \frac{\sum_{j \in J_{rmt}} p_{jrmt} \cdot s_{ijrmt}(\theta)}{\sum_{j \in J_{rmt}} s_{ijrmt}(\theta)} \right)$$

(10)

where $N_b$ is the number of income draws falling into bin $b$. Finally, as in Petrin (2002), we match the observed purchase probabilities and purchase prices to the simulated ones as our micro moments.\(^ {24}\)

5.3 Implementation details

Product set. We already restrict attention to the broader whiskey category. To further simplify our setup, we restrict attention to liquor products with the size of 750ml, thus focusing on 176 products (out of 276 from grocery retailers) that take 63.6% overall grocery liquor revenue. Focusing on one size alleviates the necessity of having to model non-IIA substitution between sizes. Miravete et al. (2017) characterize substitution between different categories (e.g. whiskey or Vodka).

\(^{23}\)To be exact, we construct this probability at the household-retailer-month level.

\(^{24}\)Conlon and Rao (2015) focus on the slope of liquor-purchase probability or purchase prices on household income, and impose that the simulated draws exactly satisfy these slopes in estimation.
and sizes with random coefficients on these characteristics.  

**Aggregation across stores and weeks.** The original data is on the level of product-retailer-store-week. We aggregate the data to product-retailer-market-month (market defined as 3 digit zip code) for two reasons. First, liquor is a slow-moving product and there are often weeks where products have 1 or 2 unit sales. It is entirely plausible that some products have zero sales in some stores for considerable number of periods. In these cases, a random coefficient logit model is not well-defined because \( \log(s_{jrm}) = \log(0) \). Aggregating the sales will considerably alleviate the problem of zero shares. Second, while there are little price variations across stores within a market, we do average over variations in prices across weeks within a month. However, because liquor is a storable product, promotional elasticities could reflect forward purchasing rather than regular price elasticities used for setting long-run price levels (Hendel and Nevo, 2006). Using monthly data helps us to focus on long-run rather than short-run price response.

**Fixed effects.** In implementation, we control for all product-retailer and retailer-market fixed effects, instead of the full product-retailer-market fixed effects. Controlling for too many fixed effects will eliminate much statistical power and risk overfitting the data. Likewise, we include retailer-year fixed effects and common year-month fixed effects. In total, we have 412 product-retailer and retailer-market fixed effects and 79 retailer-year and year-month fixed effects (relative to 174,299 observations).

**Market size definition.** We measure market size in the follow way. We take the population in the market, multiply it by the share of total grocery expenditure among the set of focal retailers in the given market-month, and multiply this by 2 to allow each person to purchase at most 2 bottles of liquor a month. With the above definition, the median market share is 0.00009. The median outside good share is 0.985 and the minimum outside good share is 0.812.

---

25 As a minor point, we also restrict the set of inside good to products that have sold at least 2,500 bottles in total, and have prices below $80. This selection criteria leaves us with 98 products but this set of products occupy 98.7% of the total revenue (98.4% of sample size after the previous sample reduction).

26 We linearly interpolate the each pair of annual population levels to obtain the monthly population levels.
**Estimation and inference.** We estimate model parameters $\theta$ via iterative generalized methods of moment (GMM). We first stack all moments denoted $g(\theta) = \begin{pmatrix} g_1(\theta) \\ g_2(\theta) \end{pmatrix}$, where $g_1$ represents moments from the aggregate data

$$
\mathbb{E}[g_1(\theta)] = \mathbb{E}[\hat{\xi}(\theta) \cdot z] = 0
$$

(11)

and $g_2$ moments from micro data

$$
\mathbb{E}[g_2(\theta)] = \mathbb{E}\left[\begin{array}{c}
\hat{s}^b(\theta) - \bar{s}^b \\
\hat{p}^b(\theta) - \bar{p}^b
\end{array}\right] = 0
$$

(12)

and $\hat{s}^b$ and $\hat{p}^b$ come from the household panel. The GMM minimizes the quadratic function of moments given the weighting matrix $W$:

$$
\mathbb{E}[g(\theta)]^T \cdot W \cdot \mathbb{E}[g(\theta)].
$$

(13)

We estimate $\theta$ with a two-step GMM algorithm. We start with the identity matrix as the initial value of $W$, and estimate $\theta$ by minimizing (13). Then, we take the previous estimate of $\hat{\theta}$ to compute $\hat{W} = \mathbb{E}\left[g(\hat{\theta}) \cdot g(\hat{\theta})^T\right]$. Finally, we use the updated $\hat{W}$ to estimate $\theta$. We find no visible difference between the one-step and two-step GMM estimates.

Following Hansen (1982) and Petrin (2002), we compute the asymptotic variance-covariance matrix of the parameters,

$$
V = (\Gamma^T)^{-1} \cdot (\Gamma^T \Gamma^T)^{-1} \cdot (\Gamma^T)^{-1}
$$

(14)

where $\Gamma$ is the Jacobian matrix $\frac{\partial g(\theta)}{\partial \theta}$. Unlike Petrin (2002), the upper off-diagonal of the Jacobian is non-zero because the aggregate moments $g_1(\theta)$ are informative about the random coefficients (due to “BLP” instruments).
5.4 Estimation results

Table 2 reports parameter estimates for the mean and standard deviation of price coefficients. We control for product-retailer, retailer-market, retailer-year and year-month fixed effects but these are not directly reported in the table. We also estimate the model without household-level coefficients and without micro moments as in Berry (1994).

In the random coefficient logit model (“main spec.”), we find considerable heterogeneity in price sensitivities and in category utility (i.e. the intercept). The 5th percentile price sensitivity is -0.432 and the 95th percentile is -0.194 – the former is more than twice as large in magnitude as the latter. Part of this heterogeneity is driven by income, reflecting that high-income consumers are less price-sensitive. Consistent with Conlon and Rao (2015), we also find that high-income consumers derive lower utility from the liquor category despite being less price sensitive.

We measure the in-sample fit of the model using the R-squared for the mean utility projection, which is inverted from applying the Berry et al. (1995) contraction mapping on observed market shares given the nonlinear coefficients. We find that the model fits data well, explaining 86.0% of the data variation.

We use model estimates to recover that $\xi_{jrmt}$ is auto-correlated. When imposing an AR(1) structure, i.e.

$$\xi_{jrmt} = \rho \xi_{jrmt-1} + \iota_{jrmt},$$

we find that $\hat{\rho} = 0.637$ (standard error = 0.002).

5.4.1 Implied elasticities and markups

We next compute implied elasticities using our demand estimates. In table 3, we present the own- and cross- elasticity matrix for six products sold by retailer 32 in June, 2016, the last month where we assume prices are optimal. We find that elasticities are increasing in magnitude with price: the implied price elasticities of these example products range between -2.18 and -4.64.

\footnote{An alternative is to estimate $\rho$ simultaneously with other structural parameters, as in Doraszelski et al. (2018).}
Table 2: Demand parameter estimates

<table>
<thead>
<tr>
<th></th>
<th>Main spec.</th>
<th>Berry (1994)</th>
<th>price (first stage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>price ($\alpha_0$)</td>
<td>-0.648</td>
<td>-0.181</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>price × log (income) ($\alpha_1$)</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>std. dev. of price coef. ($\sigma_v$)</td>
<td>0.064</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intercept × log (income) ($\gamma_1$)</td>
<td>-0.137</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>feature or display</td>
<td>0.136</td>
<td>0.117</td>
<td>-0.281</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.013)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>missing feature/display</td>
<td>0.199</td>
<td>0.195</td>
<td>-0.095</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.008)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>average price in other states</td>
<td></td>
<td></td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td>number of products</td>
<td></td>
<td></td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>product-retailer and retailer-market FE ($\delta_{jrm}$)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>retailer-year and year-month FE ($\lambda_{rt}$)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Number of observations</td>
<td>174,299</td>
<td>174,299</td>
<td>174,299</td>
</tr>
<tr>
<td>R-squared (linear part)</td>
<td>0.860</td>
<td>0.841</td>
<td>0.972</td>
</tr>
</tbody>
</table>

Notes: This table reports parameter estimates of the demand side. The first column reports estimates and standard error of the main specification. The second column reports estimates of a Berry (1994) logit model. The third column reports the first stage for price in the Berry (1994) logit model. The F-statistics for the two excluded instruments is 248.
Table 3: Example of implied elasticities and markups

<table>
<thead>
<tr>
<th>Product</th>
<th>Price</th>
<th>% Margin</th>
<th>Elasticity of: 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.01</td>
<td>0.393</td>
<td>-2.183</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>2</td>
<td>11.50</td>
<td>0.302</td>
<td>0.001</td>
<td>-2.868</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>14.53</td>
<td>0.260</td>
<td>0.009</td>
<td>0.011</td>
<td>-3.350</td>
<td>0.015</td>
<td>0.017</td>
<td>0.018</td>
</tr>
<tr>
<td>4</td>
<td>21.29</td>
<td>0.219</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>-4.095</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>5</td>
<td>26.76</td>
<td>0.202</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0.006</td>
<td>-4.507</td>
<td>0.008</td>
</tr>
<tr>
<td>6</td>
<td>29.08</td>
<td>0.198</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
<td>0.004</td>
<td>0.005</td>
<td>-4.640</td>
</tr>
</tbody>
</table>

Notes: Elasticity and implied markup for six products (these products are picked because of the differences in prices), sold by retailer 32 in June, 2016. The elasticity table reads: 10% decrease in price of product 1 will increase its sales by 21.83% and decrease the sales of product 2 by 0.04%.

Across all retailers and products, we find that the average elasticity at observed prices is -3.94. For products below $15 in average price, the average elasticity is -2.65. For products above $15, the elasticity is -4.65. These numbers are very close to Miravete et al. (2017) who find elasticities in the Pennsylvania liquor market to be -2.9 for cheap products and -4.9 for expensive products. Intuitively, this finding comes from the fact that consumers who are less sensitive to price (and thus are the main customers for high-end liquor) value the liquor category lower as a whole and thus have limited willingness to pay. Thus, retailers have lower percentage margins for high-price products.

Based on those elasticities, we compute the implied margin as a share of price if retailers set prices with full information about the demand parameters (the details of which will be shown in the next section). In the example below, for product 3, the retailer prices it at $14.53 and gains 26.0% of the price as gross margin.

5.4.2 Does demand change over time?

A potential concern in this context is that, if consumer demand changes over time after the privatization, the researcher would then need to explicitly model these changes on the demand side and the retailers might instead by learning about demand as moving (as opposed to stable) objects.

Based on our estimates, we investigate whether demand can be well-approximated by a stable function. We first examine whether the estimated time fixed effects imply that demand evolves
over time. The left-hand side panel of Figure 6 reports the average implied time fixed effects by half-year windows. We construct these time fixed effects as estimated retailer-time fixed effects $\lambda_{rt}$, having first taken revenue-weighted average across retailers, and then normalized to the level in 2012. We find that the time effects are small: the maximum changes implied by these fixed effects are 0.14 in mean utility terms, which, at the average price coefficient, implies a $0.22$ change in willingness to pay. These changes are negligible compared to the amount of price variation in the data, suggesting that demand at the aggregate is stable and consumers are unlikely to learn about the existence of the category (or many products) over time.

![Figure 6: Time trend (left) and model fit (right)](image)

**Notes:** Left panel: the time effects are retailer-time intercepts averaged across retailers and across months within half-year time windows, and normalizing to the level of 2012. Right panel: Decomposition of variance in the predicted mean utility into observables (price and promotion), product-retailer and retailer-market FE, retailer-year and year-month FE, and the error term $\xi_{jrt}$. The decomposition exercise is done by half-year windows. Estimates come from the main specification presented in Table 2.

One additional concern is that demand might change beyond what the model captures. We now examine model fit by half-year time windows. As shown on the right-hand side panel of Figure 6, the model explains the data well throughout the sample: R-squared for the mean utility is at
least 0.8 and remains stable throughout 2012 to the first half of 2016. These results suggest that demand can be well-approximated by a stable function, implying that we can indeed think of firm learning as tracking a stable object. We also show in Appendix Figure 20 that the model fits variations in market shares well for individual products.

5.5 Wholesale price (marginal costs for the retailers)

5.5.1 State-level uniform pricing with full information

In this section, we outline a price-setting model where the retailer sets prices for all of its products given full demand information. In Section 4, we found that retailer learning occurred mainly in the sample period before 2015, suggesting that we can impose full-information optimal pricing in the latter part of the data. Our main estimates for marginal costs are based on the assumption that retailers set prices optimally in Year 2016. Specifically, we note that prices decrease during late 2016 for reasons outside of the model. Therefore, we only impose the first-order conditions to prices in the first half of 2016. In addition, we allow for the possibility that wholesale prices might change over time, and use Oregon wholesale price data to recover the rate of price changes by sub-category.

Retailer \( r \) sets prices for its products with the restriction that the price must be uniform for each product across all markets in the state (we omit the state subscript). When setting prices, the retailer knows demand primitives \( \mathcal{D}_r = \{ \{ \delta_{jrm}, \lambda_{rt} \} \}_{j,m,t}, \alpha, \beta, \gamma \). In addition, we assume that the retailer does not know realized demand shock \( \xi_{jrm} \), but takes the projected demand shock from its one-month lagged value, \( \hat{\xi}_{jrm} = \hat{\rho} \xi_{jrm-1} \), which we estimate after demand estimation.

Denote \( \hat{s}_{jrm} \) the implied market shares when demand shock innovations are set to \( \hat{t}_{jrm} = 0 \).
The retailer as a multi-product and multi-market monopolist maximizes the total profit

\[
\max_{p_{rt}} \sum_{j \in J_{rt}} \sum_{m \in M_r} \left( (1 - f) p_{jrt} - c_{jrt} \right) \cdot \tilde{s}_{jrm} (p_{rt}) \cdot h_{rmt}.
\] (16)

where \( M_r \) is the set of markets the retailer operates in, \( h_{rmt} \) is the exogenous market size of market \( m \) for retailer \( r \) in month \( t \) (which is local population times the retailer’s share of grocery revenue share in the market), \( c_{jrt} \) is the wholesale price (marginal cost) for product \( j \) in \( t \). \( f = 0.17 \) is the share of gross revenue levied by the state.\(^{30}\) This profit maximization problem leads to the first-order condition such that for all \( j \),

\[
\sum_m (1 - f) \tilde{s}_{jrm} h_{rmt} + \sum_{j'} \sum_m \left( (1 - f) p_{j'rt} - c_{j'rt} \right) \frac{\partial \tilde{s}_{j'rt} (p_{rt})}{\partial p_{jrt}} h_{rmt} = 0
\] (17)

Using matrix notation, and inverting prices, reveals that:

\[
p_{rt} = \frac{c_{rt}}{1 - f} - (\Delta_{rmt})^{-1} \sum_m \tilde{s}_{rmt} (p_{rt}) \cdot h_{rmt}
\] (18)

where the \( j, j' \) th element of \( \Delta_{rmt} \) is \( \frac{\partial (\sum_m \tilde{s}_{jrm} h_{rmt})}{\partial p_{j'rt}} \).

We calculate implied markup \((\Delta_{rmt})^{-1} \tilde{s}_{rmt}\) over “effective cost” \( c_{rt} / (1 - f) \) based on demand estimates, and use them to compute retailer margin as \((1 - f) (\Delta_{rmt})^{-1} \tilde{s}_{rmt}\) divided by price. Examples of these margins for six products are shown in Table 3.

5.5.2 Wholesale price estimates

We calculate implied wholesale price for product-retailer pairs, \( c_{jrt} \), from the optimality conditions Eq. (17),\(^{31}\) which we impose on the sample of 2016. To estimate wholesale price in periods other

\(^{30}\)We examine various press articles for direct comparisons between list prices and total prices with tax, and find that both the excise tax and sales tax are excluded from the list price. See, e.g., https://www.seattletimes.com/seattle-news/the-day-liquor-went-private-and-prices-stumped-the-public/ and http://thesunbreak.com/2013/03/14/tracking-liquor-prices-in-liquor-stores-large-and-small/.

\(^{31}\)With data aggregated to the product-retailer-month level.
than 2016, we assume that
\[
\log(c_{jrt}) = \bar{c}_{jr} + \tau_{k(j)y(t)} + \omega_{jrt},
\]  
(19)

This is to say, the log wholesale price can be decomposed into a term that varies across products and retailers but is constant over time, and a term that varies across categories \((k)\) and years \((y)\) but is common across products. In addition, \(\omega_{jrt}\) are mean-zero cost shocks. We normalize \(\tau_{k(j)y(t)} = 0\) for Year 2016 for all categories and estimate \(\bar{c}_{jr}\) using our main data from Washington. Separately, we take the implied wholesale prices in Oregon where we observe both prices and markups, and estimate how wholesale prices change year-over-year and differently for different categories. Assuming that wholesale prices in Washington change proportionally in the same way as in Oregon, the Oregon estimates help pin down Washington wholesale prices in years other than 2016.

Figure 7 examines the distribution of estimated costs \(\bar{c}_{jr}\) (for Year 2016). We find that these costs have large dispersion across different products, but such dispersion is intuitive given the sharp vertical differentiation in the category. In Year 2016, the 5th percentile of costs (among product-retailer pairs) is $4.09, the median is $12.23, and the 95th percentile is $31.38.

Our estimates of category-specific price trend from Oregon indicate that liquor prices generally go up over time but the trend is different for different imported products. Compared to 2016, the prices of domestic products (mostly bourbon, rye, and some domestic whiskies) are 5.1% lower in 2012 in Oregon. This difference is mild and can be attributed to inflation. In addition, relative to 2016, Scotch are 5.2% less expensive in 2012 (very close to the general time trend), Irish whiskies are 5.8% less expensive, and Canadian whiskies are 2.0% more expensive. These differences can potentially be attributed to changes in the foreign exchange rate in our sample period: in particular, US Dollars to Canadian Dollars exchange rate moved from 1:1 in 2012 to 0.75:1 in 2016.

To check for robustness of our implied wholesale prices, we also compare our estimates to observed wholesale prices in the pre-privatization period for a set of products. Prior to mid-2012, the Washington State applies a fixed, 51.9% markup above the wholesale price for all products.
We back out the wholesale price from published retail prices using data between March to May 2012, hand-match 65 products with size 750ml that are offered both in the pre- and post- periods, and compare the two wholesale prices between these products. Although it is not guaranteed that the wholesale prices stay unchanged after a change in the retail market structure, we still find the model-implied wholesale prices to be very close to those faced by the state. The mean wholesale price is $13.8 in the privatized market (median: $12.4), similar to the $12.7 in the pre-privatization era (median: $12.1). In addition, across products, the model-implied- and state- wholesale prices have a correlation coefficient of 0.99. These results provide external validity that the model does a good job recovering marginal costs for the retailers.

6 How much, and in what ways, do retailers learn?

In this section, we contrast observed prices to model-implied optimal prices and descriptively study how much, and in what ways, retailers learn about demand and improve prices. We focus
our supply-side analysis on three large grocery retail chains that capture 80.7% of the liquor market (in revenue shares). For the three excluded retailers, two of them are drug stores (Retailer 4901 and 4904), and one of them is a grocery chain (Retailer 9) with a small share of the Washington liquor market. We exclude this grocery chain because it closed a large fraction of its stores during our sample period.

We start by using the model to simulate a counterfactual scenario in which retailers have full information about demand primitives and set prices optimally. This counterfactual provides the normative comparison by which we evaluate the potential for, and the pace of, learning. We find large gaps between observed prices and model-implied optimal prices in the initial period, suggesting retailers, especially those without prior experience in other liquor markets, make large initial mistakes. These large price gaps suggest that the returns to having full information at the start of the market amount to up to 13% of profit for inexperienced retailers.

Next, we descriptively study what retailers might learn about, and find that the observed pricing patterns are suggestive of learning about both the price-sensitivity of the customers and individual product quality (i.e. utility intercepts). Finally, we explore other possible drivers of learning and find that, beyond information from prior experience, retailers also learn from prices set in other states before entering the Washington market.

### 6.1 Full-information optimal prices

We first use the structural model to simulate the counterfactual where retailers have full information and set prices optimally. This counterfactual provides a benchmark with which to compare actual decisions. Such a model allows us to evaluate how close to perfect-information behavior the retailers’ initial behavior is, a signal of their initial knowledge position, and how long it takes them to set prices that are close to the perfect information optimal levels.

We compute model-implied optimal prices using estimates of the demand and marginal costs

\[
\hat{c}_{jrt} = \exp \left( \tilde{c}_{jr} + \tau_{k(j)y(t)} \right),
\] (20)
setting the cost shocks $\omega_{jr}$ to zero. We plug in the demand and cost primitives to the corresponding first-order conditions, as defined in Section 5.5.1, and solve for optimal prices for all retailers in all months in Washington.\textsuperscript{32} Specifically, we jointly solve for $p^*_{rt}$ as the implied full-information optimal price vector, defined as a solution of the Bertrand-Nash price equation

$$p^*_{rt} = \frac{c_r}{1 - f} - (\Delta_{rmt})^{-1} \sum_m \tilde{s}_{rmt}(p^*_{rt}) \cdot h_{rmt}, \quad (21)$$

where as before, $\tilde{s}_{rmt}$ is the vector of market shares as a function of price, with $\tilde{\xi}_{jrmt}$ set to $\hat{\rho}\tilde{\xi}_{jrmt-1}$, the predictable part of the demand shock. Equation (21) needs to be solved as a fixed point, jointly for all products in each retailer-month.\textsuperscript{33} After obtaining $p^*_{rt}$, we compute the implied total profit for retailer $r$ in year-month $m$ across all products and markets:

$$\pi^*_{rt} = \sum_j \sum_m \tilde{s}_{jrmt}(p^*_{rt}) ((1 - f) p^*_{jrt} - \bar{c}_{jr}) \cdot h_{rmt}. \quad (22)$$

We contrast these “optimal full-information profits” against the profit evaluated at the observed prices.

### 6.2 How important is learning across experienced and inexperienced retailers?

We first measure the scope of learning by contrasting observed prices to the optimal prices implied by the model. We construct a measure “%price gap” capturing the relative difference between the optimal and the observed prices, or

$$\text{%price gap} = \frac{p^*_{jrt} - p_{jrt}}{p_{jrt}}. \quad (23)$$

\textsuperscript{32}Unlike Chintagunta et al. (2003), we do not simulate optimal local prices, since we observe that retailers set prices at the state level, see section 3.3.1.

\textsuperscript{33}We use an optimizer to directly maximize profit, and find that iterating on Equation (21) gives profit-maximizing prices.
We also measure the relative difference in the retailers’ total profit from the full-information optimum, $\sum_r (\pi^*_r - \pi_{rt}) / \sum_r \pi_{rt}$.

One might expect retailers’ prior experience selling liquor in other markets to play an important role in shaping their initial prices and learning rates in this market. Retailers that have never sold liquor in other states might be less informed about demand. Consequently, these inexperienced sellers might set prices further away from the full-information optimum and have more to learn. In our sample, Retailer 158 and 182 are “local” retailers which operate only in Washington and some other liquor-control states. As a result, they operated in a deregulated liquor market for the first time when Washington was privatized. In contrast, Retailer 32 has previous experience selling liquor in other states that spans many years. We contrast the learning behavior of the two groups of retailers.

The upper panels of Figure 8 present %price gaps over time, separately by inexperienced and experienced retailers. The bottom panels present the relative differences in profit. We find that inexperienced retailers set prices higher than the full-information optimum. Averaged across the first half-year period, their price levels are 8.7% higher than the optimal level in the median. In addition, there are large dispersions in the price gaps across products: for example, the inter-quartile range of the %price gaps is 18.5% (quartiles are [-1.0%, 17.5%]). As a result, total variable profit at observed prices is 13.0% lower than the variable profit at full-information optimal prices – a dramatic difference from the lack of information.

We find that soon after the privatization, these retailers set prices closer to the full-information optimum. In the second half-year window since the privatization, the median price gap reduces to 3.0% and the inter-quartile range reduces to 14.8%. These changes convert to a 5.9% decrease in the profit gap (to 7.1%) in the first half-year period, implying large gains from initial learning for these inexperienced retailers. In longer horizons, prices continue to improve and the profit gap further decreases by 6.5% in the next 3 years, to only 0.4% in the first half of 2016. We conclude that inexperienced retailers make large initial mistakes but are able to learn and correct these mistakes.
Figure 8: Percentage price and profit gap by experienced and inexperienced firms

Notes: Top panels: distribution of the percentage price gap (positive or negative) between observed prices and full-information optimal prices, where the distribution is across products and retailers. Bottom panels: percent total profit gap between observed prices and full-information optimum. Left panels: inexperienced retailers. Right panels: the experienced retailer.
In contrast, Retailer 32 has operated in the liquor market in other states for many years. In the initial half-year window, this experienced retailer sets prices that are 5.2% too high compared to the optimal prices in the median and 12.3% in the inter-quartile range, implying a profit gap of 6.4%. Over time, the retailer adjusts prices closer to the full-information optimum (both reflected in the median and the IQR) and improves profit by 3.8% within the sample.\footnote{The price level of this retailer appears to be lower than the full-information optimum in 2015. We examine the pricing strategy in detail and find that Retailer 32 practices promotions in this period and high-low prices are not fully captured by the model.} We find that, while the experienced retailer makes fewer (or smaller) mistakes initially compared to the inexperienced retailer, there is still a meaningful improvement over time consistent with learning.

6.3 What might retailers learn about?

We next ask what actual features of demand retailers learn about. For a retailer new to the liquor market, it might not know either the quality of individual products (as retailer-product intercepts $\delta_{jr}$) or the distribution of price sensitivities (as $\alpha_i$ in the model). If the retailer does not know about the distribution of $\alpha_i$, it will initially set either too high or too low prices for all products. In this case, the \%price gaps will be on one side of zero (but potentially with different magnitude).

The pricing mistakes and learning pattern will be different if the retailer is informed about the distribution of $\alpha_i$ but not about individual product quality $\delta_{jr}$.\footnote{This argument also assumes that 1) the retailer is informed about the distribution of $\delta_{jr}$ (so that it knows the mean and variance of product quality) and 2) cross-elasticities are so small that it can be ignored, which we confirm in our demand estimates.} For products whose qualities are initially over-estimated, the retailer will initially over-price and then adjust prices downward as it learns; for products whose quality are initially under-estimated, prices will initially be too low and then increase over time. This case implies that the gap between observed and optimal prices will initially appear “scattered” (with mean zero) and shrink as the retailer learns.

We first re-visit the upper panels of Figure 8 where we plot the distribution of \%price gaps over time and by inexperienced and experienced retailers. While we have shown that inexperienced retailers make larger initial mistakes, suggesting that they are more constrained by information,
both sets of retailers exhibit similar qualitative patterns. First, overall price levels are too high. This pattern is consistent with the conjecture that these retailers have incorrect beliefs about the consumer price sensitivity distribution. Second, many products are initially priced far away (higher or lower) from the optimal price and the distribution of price gaps shrink over time. In particular, while the majority of products are initially over-priced, more than 25% of products are under-priced suggesting that the retailer under-estimated the demand for some products. This pattern suggests that the retailers also have limited information about product quality ($\delta_{jr}$) when the market starts.

Why do retailers need to learn about consumer price sensitivities? One possible explanation is that, as our demand estimates imply, higher-income customers are more averse to liquor. As a result, liquor customers are lower-income grocery shoppers and many of them cannot afford high-priced products. As we show in Table 3, the optimal percent profit margin should be inversely related to the price level. Whereas an experienced retailer might be aware of this feature of the market, a retailer who has never operated in any liquor market might have the tendency to charge flat percent margin or even higher percent margin for high-quality products. Another explanation is that Washington levies a stiff 17% licensing fee, which increases the retailer’s effective costs and forces them to set prices in a region of demand where they are less informed about the local price elasticities.

To descriptively examine whether retailers initially have imperfect knowledge of, and learn about, liquor-shopper demographics and the “correct” tax pass-through, we summarize the realized percent margins (i.e. price minus cost divided by price) as a function of model-implied optimal prices. Optimal prices are higher for products with higher intercepts and can be interpreted as a proxy for product quality. Figure 9 show the results separately for inexperienced (upper panel) and experienced (lower panel) retailers and for three cross-sections of the data.

For inexperienced retailers in the first quarter after the privatization, we find that while low-optimal-price (i.e. low-quality) products have on average the correct margin, the margins for mid- and high-price products are much higher than implied by the model. This profile of initial mis-

---

36 Miravete et al. (2017) show a similar pattern in Pennsylvania.
37 We do not directly examine intercepts because they are not comparable across the two retailers.
Figure 9: Percent-margin and optimal prices across products: inexperienced (top) and experienced (bottom) retailers

Notes: the x-axis is log optimal price implied by the model, the y-axis is percent observed (solid, and circles) and optimal margin (dash), defined as price minus marginal costs as a fraction of price. Circles are product-retailer pairs, whereas the lines are local polynomial fit. We omit the individual points for optimal markup to keep the figures clean, but note that the optimal markups are close to the dashed line.
takes is consistent with the hypothesis that these retailers have limited knowledge about customer demographics: initial prices are suggestive of the retailers trying to sell to customers with higher income and thus over-price the mid-tier and high-end products. In a year, the (average) observed margin profile “rotates” and matches the shape of the optimal profile, despite that the observed margin for all price-tier products is still a few percentage points higher than the optimum. The observed margin profile converged to the optimum in another year.

In contrast, the learning patterns for the experienced retailer seem to be different. We find that initially, the (average) margin profile is about 5 percentage points higher than the optimum, consistently across low-, mid- and high- price tier products. This pattern is consistent with the conjecture that, while the retailer understands that percent-margin should be lower for high-quality products, it does not know the correct amount of taxes to pass-through to the consumers under a unique tax (and fee) regime in Washington. Over time, however, the retailer is able to adjust prices to be closer to the optimum.\textsuperscript{38}

6.4 Do retailers learn from others?

In this section, we provide additional descriptive evidence about what sources the retailers obtain information from. We show in Section 4.1 that retailers learn from their own quantity shocks, and section 6.2 that prior experience in liquor markets transfer to Washington and play an important role in determining the initial prices set in the new market. In this section, we focus on whether retailers learn from information from other states or from prior experience by the Washington state-owned liquor chain.

Indeed, (privatized) markets in other states might provide information about consumer demand in Washington. This information might be useful because, across states, the market position of individual products are correlated and liquor shopper demographics are similar.\textsuperscript{39} If Washington retailers learn from other states, we should expect that their initial pricing mistakes are smaller

\textsuperscript{38}We note that the lower observed prices in 2014Q3 are associated with a period of frequent discounts.

\textsuperscript{39}We find that, for a fixed set of products, the correlation coefficient in sales rank in Washington and in all other states (within the Nielsen data) is 0.71.
for products that are widely carried in other states, relative to products that are not available, or not widely carried, in other states. On the contrary, Washington retailers might not learn from the state-owned chain as much, because the state-owned chain charged a fixed markup (i.e. did not practice demand-driven pricing).

To test these conjectures, we take the median split of the average sales volume in all other states within the Nielsen data, and split the set of products in our sample into two groups: “popular” and “unpopular” in other states. We also group products into those “carried by” and “not carried by” the Washington state-owned chain as indicated by the state’s price list. These two groups create a 2x2 design. Our conjectures suggest that initial mistakes are smaller for products popular (versus not popular) in other states, but are no different for products carried (versus not carried) by the Washington state-owned chain.

Figure 10 presents initial pricing mistakes measured by the %price gaps relative to the full-information optimum. We focus on the first quarter and on retailers without own experience in other markets. We find that retailers make smaller mistakes for products that are popular in other states (likely, they are widely carried) compared to those unpopular in other states. This finding is consistent with the conjecture that retailers learn from prices set by others in similar markets, and use this information to set initial prices in the new market. On the contrary, we do not find pricing to be closer to optimum for products carried by the state firm. This finding is consistent with the conjecture that the state-owned chain does not have knowledge about demand that can “pass on” to the new retailers.

7 Conclusion

How do new entrants learn about demand and update their pricing strategies? We study the changes in retailers’ pricing policies in the Washington State liquor market, where the privatization of liquor sales leads to existing grocery chains entering the liquor market for the first time. This context is

---

40 We record sales volume of zero for products not available in other states.
41 The differences are most pronounced for products not carried by the Washington state-owned chain.
ideal to study our question because the policy change exogenously created new and (potentially) inexperienced sellers in an existing market where demand is stable over time so that retailers’ learning to set prices can be isolated from many possible alternative stories.

Shortly after the privatization, retail prices show large changes. We document two pieces of descriptive evidence showing that these price changes are consistent with retailers learning about demand. First, prices respond to lagged demand shocks and the rate of response declines over time. Second, prices are more correlated with sales as retailers’ experience in the market increases. We then estimate a structural model of demand and costs in order to simulate a counterfactual where retailers have full-information and set optimal prices based on this information. Comparing to this full-information counterfactual, we present three conclusions. First, retailers do learn about demand over time and prices become more similar to the optimal full-information levels. Second, early in the market, prices are set to sub-optimal levels compared to the (theoretical) perfect-information case. In particular, the lack of optimality corresponds to as much as a 13% loss in gross profit for retailers who are first-time liquor sellers. Third, the contrast between observed

Figure 10: Initial %price gaps by product groups

Notes: Focus on %price gaps in the first quarter by the inexperienced retailers. Box plot presents 10th, 25th, 50th, 75th, and 90th percentiles. Popular product defined as average sales volume ranked above median.
and optimal prices suggest that learning follow intuitive patterns. We find that the initial prices set by the retailers suggest they lack initial knowledge about 1) demographics of liquor customers, 2) price elasticity of demand under the high Washington taxes, and 3) individual product demand in Washington. We also show that retailers not only learn from own experience but also from the behavior of retailers in other states.
References

Adams, Brian and Kevin R Williams (2017), ‘Zone pricing in retail oligopoly’.


Arcidiacono, Peter, Paul B Ellickson, Carl F Mela and John D Singleton (2016), ‘The competitive effects of entry: Evidence from supercenter expansion’.


Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron S Jarmin, Megha Patnaik, Itay Saporta-Eksten and John Van Reenen (2017), ‘What drives differences in management?’.


Conlon, Christopher T and Nirupama Rao (2015), ‘The price of liquor is too damn high: Alcohol taxation and market structure’.


DellaVigna, Stefano and Matthew Gentzkow (2017), ‘Uniform pricing in us retail chains’.


Illanes, Gaston and Sarah Moshary (2017), ‘What is the marginal effect of entry? evidence from a natural experiment in liquor licensure’.

Jeon, Jihye (2017), ‘Learning and investment under demand uncertainty in container shipping’.

Li, Yang, Brett R Gordon and Oded Netzer (2017), ‘An empirical study of national vs. local pricing by chain stores under competition’.


Miravete, Eugenio J, Katja Seim and Jeff Thurk (2017), ‘One markup to rule them all: Taxation by liquor pricing regulation’.


Seo, Boyoung (2016), ‘Firm scope and the value of one-stop shopping in washington state’s deregulated liquor market’.


### A Evidence for uniform pricing

#### A.1 Example

Figure 11 follows DellaVigna and Gentzkow (2017) and visualizes price variations over time and across markets for the top-selling product across all six retailers, Fireball Canadian whiskey. We take the log price for the product in each 3-digit-zipcode in each month, normalize it by its mean within each retailer, and plot the normalized log price deviations in color grids of 0.1. Dark greys show markets or months with higher prices, and light greys show markets or months with lower prices.

For this product, we find that prices change over time in correlated ways across markets. Prices for Retailer 4901 is the most uniform as there are no visible differences across markets at the same time. Even for Retailer 9 or 32 where there are some price variations across markets in a given month, price increases or discounts are usually synchronized across most markets. The example shows that prices are close to uniform in Washington State and suggests that each retail chain makes pricing decisions at the state level. We check and find similar patterns for other products, and show formal evidence of state-level pricing in Section A.2.

#### A.2 Test of uniform pricing

We formally examine the degree of price dispersion for given products across stores and chains, and show that prices seem to be set on the chain level. To examine the cross-sectional aspects, we first estimate simple specifications of log prices on various levels of fixed effects (FEs):

\[
\log \left( p_{jrmt} \right) = \alpha_X + \epsilon_{jrm}, \tag{24}
\]

where, with an abuse of notation, the subscripts of \( \alpha_X \) takes product level (i.e. \( \alpha_j \)), product-retailer level (\( \alpha_{jr} \)) or product-retailer-market level (\( \alpha_{jrm} \)) respectively. Because liquor products are naturally vertically differentiated, we do not interpret differences in the price levels across
Figure 11: Price variations over time and across markets: example

Notes: This figure follows DellaVigna and Gentzkow (2017) and visualizes price variations over time and across markets for Fireball Canadian whiskey across all six retailers. We take log price for the product in each 3-digit zipcode in each month, normalize by the mean within the retailer, and plot the deviation in 0.1-grid. White represents that the product (or the retailer) is not present in the market.
products. Instead, conditional on the estimated product FEs $\hat{\alpha}_j$, we focus on whether additional FEs at product-retailer or product-retailer-market levels help explain the “left-over” variations in log prices. Specifically, to examine the extent to which product-retailer FEs explain price variations, we calculate both the incremental amount of price variations explained by product-retailer FEs, $SSR^{jr} - SSR^j$, and the total variations not explained by the product FEs, $SST - SSR^j$. The ratio between the two,

$$\frac{SSR^{jr} - SSR^j}{SST - SSR^j},$$

measures the amount of price variations explained by adding product-retailer FEs on top of the product FEs. We calculate the incremental fit in the same way for the model with product-retailer-market FEs.

Relative to having product FEs only, we find that adding product-retailer level FEs explains 30.9% of the “left over” price variations. This number indicates that there are large differences in the price levels across retailers for the same product. However, adding the product-retailer-market level fixed effects only captures 1.7% (percentage points) additional price variations, suggesting that the meaningful cross-sectional price variations happen at the product-retailer level.

Next, we estimate a series of regressions with product-retailer-market FEs, but with linear time trends that vary at different levels. This is to say, we estimate a class of regressions

$$\log(p_{jrm}) = \alpha_{jrm} + \beta_X \cdot t + \epsilon_{jrm},$$

where we allow the coefficient on time, $\beta_X$, to vary at the product, product-retailer, or product-retailer-market level. Measuring incremental fit in the same way, we find that product-level time trend explains 48.7% left-over variations from the model with only product-retailer-market FEs, implying a 16.1% incremental fit than without the product-level trend. Product-retailer trends explain 6.2% more of the variation, while product-retailer-market trends only capture an additional 1.4%. This is to say, both price levels and price variations over time occur at the product-retailer level.

---

42Where $SST$ is the sum of squared of the dependent variable (as a measure of total variation), and $SSR$ is the sum of squared of the regression fit (as a measure of model-predicted variation).
Figure 12: Incremental explanatory power of co-variates to price variations

Notes: This figure reports incremental fit measures defined in (25), across log price regressions with different sets of fixed effects: product(j)-retailer(r), product-retailer-market(m), and from this version, adding product-time trend, product-retailer time trend and product-retailer-market time. The benchmark regression we use to compare fit is one with only product fixed effects.

\[ \text{B Test for cross-chain substitution} \]

In this section, we test whether demand for a given product substitutes between chains in the local market. Denote \( j \) as a product (pooled across sizes), \( r \) as a retail chain, \( z \) as a 5-digit zipcode, and \( t \) as a month. We estimate a linear model of log sales quantity on the availability of the product in the same chain and in other chains in the market:

\[
\log(q_{jrzt}) = \beta_1 n_{\text{store}_{jrzt}} + \beta_2 n_{\text{store}_{j,-r,zt}} + \delta_{jrz} + \psi_{jt} + \epsilon_{jrzt}. \tag{27}
\]

Here, \( n_{\text{store}_{jrzt}} \) is the number of stores in chain \( r \) within the 5-digit zipcode \( z \) where product \( j \) is available, and \( n_{\text{store}_{j,-r,zt}} \) is the number of stores in other chains (among the six focal chains) where \( j \) is available. Nielsen RMS data only identify store location up to 3-digit zip-
Appendix Table 4: Sales quantity on availability and promotion of other retailers

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#stores carrying the product, own chain ($\beta_1$)</td>
<td>0.2578***</td>
<td>(0.0063)</td>
</tr>
<tr>
<td>#stores carrying the product, other chains ($\beta_2$)</td>
<td>0.0002</td>
<td>(0.0012)</td>
</tr>
<tr>
<td>#share of stores on promotion, own chain</td>
<td>0.0797***</td>
<td>(0.0056)</td>
</tr>
<tr>
<td>#share of stores on promotion, other chain</td>
<td>0.0019</td>
<td>(0.0066)</td>
</tr>
<tr>
<td>product-retailer-market ($\delta_{jrm}$)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>product-year-month ($\psi_{jt}$)</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Notes: The effect of product availability and promotion, in the own chain and in other chains, on sales quantity.

code level, which is too large to measure spatial substitution. We proxy chain location using the most frequently-appearing household 5-digit zipcode among shoppers to the chain.\(^{43}\) In addition, we control for product-retailer-zipcode fixed effects and product-time fixed effects. With these controls, and also keep in mind that we focus on products that are available at the start and at the end of the sample, variations in the number of stores carrying the product comes from store entry and exit (see Figure 14) and from stores not carrying the product in a subset of weeks (likely, these are due to stockouts). In addition, we also test whether promotion in other chains in the market reduce sales in the focal chain. In a similar model as (27), we estimate the effect of share of stores on price promotion (defined as price at least 5% below the regular price), for the focal chain and for other chains.

Table 4 reports the estimation results. We find that if the product is available in one more store in the same chain, sale quantity increases by 26%. This finding is coherent with Seo (2016), in that local availability is important to demand. However, availability in one more store in other chains do not affect sales quantity in the focal chain. Similarly, our estimates on price promotion suggest that the share of stores where the product is on promotion does not affect sales quantity in the focal chain. These results suggest that liquor sales do not substitute between chains.

\(^{43}\)Our approach follows an earlier version of Illanes and Moshary (2017).
Figure 13: Correlation between observed price and average product-retailer FE in Washington

Notes: correlation coefficient between observed price and the model-recovered $\bar{\delta}_{jr} = \frac{1}{M} \sum_m \delta_{jrm}$.

C Do prices capture product-retailer demand?

As an alternative approach to Section 4.2, we now examine the price correlations with the demand intercepts using our structural estimates of the demand intercepts, focusing on the three large retailers. We consider the extent to which prices capture heterogeneity in local (WA state) tastes by calculating the correlation coefficient between price $p_{jrt}$ and the model-recovered average product-retailer fixed effect, $\bar{\delta}_{jr} = \frac{1}{M} \sum_m \delta_{jrm}$ for different periods of time. As retailers learn about the demand intercepts in Washington, their prices should increasingly reflect the $\bar{\delta}_{jr}$'s. Figure 13 shows that the correlation coefficient between $p_{jrt}$ and $\bar{\delta}_{jr}$ increases over time in Washington, consistent with our descriptive evidence in Section 4.2. These increasing correlation coefficients imply that observed prices reflect the underlying differences in demand more and more over time. Similar to the discussion in Section 4.2, the correlation between consumer preferences and prices reflects the accuracy of retailers’ beliefs about demand, and the improvement in this accuracy implies that retailers learn about demand (and improve the prices they set) over time.
D  Additional tables and figures

Figure 14: Number of stores selling grocery (solid) or liquor (dots)

Notes: Solid: number of distinct store IDs for each chain selling grocery. Dots: number of distinct store IDs selling liquor. For measures of stores selling liquor, we cross-checked with the number of license holders reported by WSLCB and find identical results. The 6 chains are chain 9 (top-left), 32 (top-right), 158, 182, 4901, 4904.
Figure 15: Total liquor sales revenue over time

Notes: Sales revenue for the focal 7 chains from the liquor category in half-year intervals. White: across all products. Blue: across core products defined in Section 3.1.

Figure 16: Price changes over time for each retailer-product (initial value = 0)

Notes: The y-axis represents the relative differences between price and initial price for a given product. Blue = final price lower than initial price. Red = final price higher than initial price.
Figure 17: Descriptive evidence for consumer behavior in the WA liquor market

Notes: These panels present additional descriptive evidence from the household panel. In the top-left, we present changes in liquor expenditure before and after the privatization. For this panel, we estimate linear regression of log liquor expenditure, \( \log(\text{expd} + 1) \), on a set of half-year dummies, for all consumers in Washington state in the household panel. The figure reports these regression coefficients. We re-defined half-years to “December to May” and “June to November” in order to align with the timing of the policy change. In the top-right, we present share of liquor expenditure in a trip to grocery store. The bottom panels are additional measures of varieties: the number of distinct chains, liquor product types, brands, and bottles of liquor per month, for trips with liquor purchases.
Figure 18: Liquor expenditure in the household panel

Notes: Liquor expenditure in the household panel, across the focal 6 retailers, and including the state store (or former state store) and other players.
Figure 19: Price ratio between Scotch/Irish whiskey versus Bourbon/US-made whiskey

Figure 20: Model fit by brand

Notes: Observed market shares in circles and model-predicted market shares in lines. The model-predicted shares use $\xi_{jm}$ draws from its empirical distribution.
Figure 21: Total profit under observed and model-implied optimal prices

Notes: The profits are sum of profit from the six grocery retailres in given half-year window.