Retailer Price Competition and Assortment Differentiation: Evidence from Entry Lotteries*

Brett Hollenbeck
Anderson School of Management
UCLA †

Sylvia Hristakeva
SC Johnson College of Business
Cornell University ‡

Kosuke Uetake
School of Management
Yale University§

February 2026

Abstract

This paper studies how local competition affects pricing and product assortment decisions. We leverage a unique circumstance in the context of the legalized cannabis industry in Washington State, where retail licenses were allocated via a lottery, which generates quasi-random variation in the number and proximity of competitors. This natural experiment helps address the endogeneity issues commonly encountered in causal analyses of competitive effects. The analyses yield three key findings: first, additional nearby competitors reduce markups, with nearly all of the effect concentrated in the first two competitors. Second, retailers facing more nearby competition differentiate their assortments. Third, this differentiation helps mitigate the intensity of price competition among retailers. In a partial equilibrium simulation, we evaluate a 10% increase in the number of licensed retailers and estimate that it would reduce markups and generate approximately $7.5 million in annual consumer savings. These findings shed light on the nature of localized competition and the implications of entry restrictions in regulated retail markets.

Keywords: Retail competition, causal inference, markups, firm differentiation, spatial competition, entry regulation, legal marijuana

*We thank Judy Chevalier, Jūra Liaukonytė, Jesse Shapiro, and conference/seminar participants at NBER IO Conference, the Triangle IO Conference, UCLA, Cornell University, UC Berkeley, Rice University, Carnegie Mellon University, and University of Toronto Rotman for helpful comments and suggestions.

†brett.hollenbeck@anderson.ucla.edu
‡sylvia.hristakeva@gmail.com
§kosuke.uetake@yale.edu

1 Introduction

Understanding how firms respond to competition is a fundamental question in industrial organization and marketing. Standard models predict that competition lowers markups and may induce firms to differentiate their offerings (Tirole (1988)), but the magnitude and shape of these effects depend on market primitives that theory alone cannot pin down. Regulatory policies and business decisions, from entry regulation and merger evaluation to pricing and product assortment strategy, depend on the nature and intensity of these competitive responses, yet credible causal evidence remains limited, especially in markets where competition is shaped by geographic proximity. This paper provides new causal evidence on how competition affects both retail prices and product assortment decisions, and examines how these two responses interact.

Evaluating these effects is empirically challenging because firms choose both where to enter and how to set prices based on local demand and competitive conditions. We exploit a natural experiment in the Washington State cannabis industry where a retail license lottery generated quasi-random variation in the number and spatial proximity of competing retailers. This variation allows us to isolate the causal impact of competition on firms' equilibrium strategies without imposing assumptions on demand or firm conduct.

We focus on three key research questions. First, how does local competition affect retailer markups? While theory suggests that competition lowers prices, the magnitude and functional form of this relationship remain open empirical questions.[1] We provide direct causal evidence on this relationship, using measures that capture both the number of competitors and their spatial proximity. This allows us to evaluate empirically how localized the competitive effect on markups is and whether there are diminishing returns from additional competitors. We also examine whether these effects differ between independent retailers and stores affiliated with a chain.

Second, how do retailers adjust their product assortments in response to competition? Confronted with a nearby competitor, a retailer may attempt to soften price competition by differentiating itself (Ginsburgh, de Palma, Papageorgiou, and Thisse (1985)), and for retailers, product assortment is the primary margin of differentiation (Dukes, Geylani, and Srinivasan (2009)). We examine whether and how firms expand or shift their product mix in response to local competition. This includes studying whether competition leads retailers to differentiate horizontally by offering different product categories to match varying consumer tastes, or vertically by supplying products of different quality and price points.

Third, how do pricing and product differentiation interact in shaping market outcomes? Firms adjust prices and assortments jointly, and these strategies may reinforce or offset each other.

[1] Theoretical frameworks have also identified mechanisms where the relationship between prices and competitive pressures could be ambiguous or even positive (Satterthwaite (1980), Stiglitz (1987), Rosenthal (1980), Armstrong and Vickers (2022), Chen and Riordan (2008), Zhu, Singh, and Dukes (2011)). Empirical counterexamples include Thomadsen (2007), which finds higher prices in duopoly than monopoly markets in fast food, and Arcidiacono, Ellickson, Mela, and Singleton (2020) which finds that Walmart Supercenter entry does not cause incumbent grocery stores to lower prices.

Whether differentiation softens or intensifies price competition is theoretically ambiguous. Standard models predict that horizontal differentiation should decrease price competition by allowing firms to carve out distinct market segments. By contrast, in settings with search costs Anderson and Renault (1999) show that prices may fall with the degree of product differentiation because more variety leads to more search, thereby intensifying price competition. We study this interaction by decomposing competitive pressure into store-level and product-level components, testing whether product overlap between nearby retailers intensifies price competition beyond store-level effects.

Our empirical context of the legalized cannabis market in Washington State provides an ideal setup to study these questions. First, we observe unusually detailed administrative data recording both retail and wholesale prices at the product level, allowing us to measure retailer markups directly. Second, we exploit a unique circumstance that created quasi-random variation in both market structure and spatial competition. In 2014, Washington State allocated retail licenses via lotteries in jurisdictions where the number of applicants exceeded the number of allocated licenses. Applicants were required to secure store locations before the lottery, and we observe the full set of applications, not just the realized winners. This allows us to characterize both the competitive environment each retailer actually faced and the competitive environment it would have faced in expectation across lottery draws.

A challenge for identification is that the pre-lottery location choices of the applicants are endogenous. Firms selected their proposed locations based on expected profitability before the lottery occurred, meaning that areas with more applicants were likely perceived as more attractive. This non-random exposure to the exogenous lottery shock introduces omitted variable bias. To address this concern, we adopt the strategy introduced by Borusyak and Hull (2023), and decompose realized competition into (i) expected competition, which is a known function of the applicant configuration and lottery probabilities, and (ii) the lottery-driven residual. Including expected competition as a control absorbs the endogenous component, so that identification comes from whether a retailer ended up with more or fewer competitors than expected. That is, our identifying variation is driven purely by random lottery outcomes. Specifically, our markup analyses regress firms' average markups at the retailer-by-product-by-quarter level on local competitive pressures, while controlling for expected competitive pressures.

We measure the degree of local competition in two ways: (i) the distance to the nearest competitor's store and (ii) the number of competitors within a set of prespecified radii. We find that distance to the closest competitor has a strong causal effect on markups. The competitive effect is highly localized: concentrated within roughly 1km and attenuating sharply beyond that distance. For example, we find that retailers whose closest competitor is within 0.2km charge markups approximately \$0.70 lower (about 9% of the mean) than retailers whose nearest competitor is more than 1.2km away. Next, we focus on the potentially nonlinear relationship between markups and the number of nearby firms. We find that almost all of the decrease in markups is caused by the first and second nearby competitors, consistent with the results in Bresnahan and Reiss (1991). Finally, we examine how competitive effects vary by retailer type. Independent stores respond more strongly to nearby entry than chains. This heterogeneity highlights the importance of market structure beyond just the number of firms.

Our second set of results examines how competition affects the size and composition of retailers' product assortments. We document that retailers facing more competitors carry larger assortments. To measure differentiation, we also compare the degree of product overlap between each retailer and its competitor in the closest geographic proximity. While we might expect that stores that are closer together would serve more similar customers and face upstream cost advantages from offering the same set of products, we find the opposite: closer competitors carry more differentiated assortments. We then decompose this differentiation by asking whether stores specialize in higher-quality or lower-priced products, or whether stores specialize in specific product categories.[2] We find stronger evidence for horizontal differentiation, with stores facing stronger competition reacting by differentiating at the category level.

Third, we examine whether firms are actually able to soften price competition through differentiation. We decompose competitive pressure into a store-level component (the number of nearby retailers) and a product-level component (the number of nearby retailers carrying the same product). This allows us to test if nearby retailers selling the same product exert more downward pressure on product-level markups than the mere presence of a competitor. Our findings reveal that product-level competition impacts markups above and beyond the overall degree of competition. For example, relative to a local monopolist, duopolists selling the same product set markups \$0.65 lower, compared to \$0.41 for duopolists selling different products. These results underscore the important role of assortment differentiation in weakening retail price competition and show that firms' pricing and differentiation strategies cannot be evaluated in isolation.

Beyond quantifying competitive responses, our study has direct implications for market design and regulatory policy, especially for markets subject to entry restrictions.[3] Entry caps create economic rents for lottery winners by restricting competition and allowing for higher markups, which effectively transfers surplus from consumers to retailers. Using our causal estimates, we conduct simulations to evaluate how relaxing entry restrictions affects market outcomes in a partial equilibrium approach, which holds quantities sold fixed. We simulate a 10% increase in the number of licensed retailers, drawing from the observed pool of applicants and their proposed store locations to account for spatial heterogeneity in competition. We estimate annual consumer savings of approximately \$7.5 million through lower prices. These effects are most pronounced in urban areas and in markets where the original lottery was highly oversubscribed underscoring both the localized

[2] Our use of product categories to capture horizontal differentiation is motivated by the literature in marketing noting that consumers view retailer positioning primarily through the stores emphasis on different category offerings (Briesch, Dillon, and Fox (2013), Krishan, Koelemeijer, and Rao (2002), Mogilner, Rudnick, and Iyengar (2008)). Our use of pricing to capture vertical positioning is also well-established in the marketing literature (Bolton and Shankar (2003), Bolton and Shankar (2004)).

[3] See Pozzi and Schivardi (2016) for an extensive overview of the types of regulations and measures used by governments to restrict retail entry as well as trends in these regulations and empirical evidence on their effects on consumers and firms. Regulated industries like the cannabis, alcohol, and tobacco industries are especially likely to see entry restrictions as we discuss in Section 7.

nature of retail competition and the inefficiency of restricting entry through license caps. While partial in scope, the exercise illustrates the trade-offs regulators face in balancing consumer welfare and firm profitability under constrained entry.

This paper contributes to the understanding of how local competitive pressures shape retailer behavior. Utilizing a unique natural experiment provided by the lottery-based allocation of retail licenses, we causally identify the impact of local competition on both the pricing and product assortment strategies, tracing out nonlinear effects along both spatial proximity and the number of competitors. Our estimates, obtained without imposing assumptions on consumer demand or firm conduct, provide direct empirical guidance on the scope and nature of local retail competition. Ultimately, this understanding is an important input into a range of analyses, including potential retail mergers, zoning-based entry regulations, and models of retail strategy that require clear guidance on the scope of local competitive effects and how both prices and product variety will likely respond.

Relationship to Literature We contribute to the empirical literature on how market structure affects retail pricing and product assortment strategies. These two dimensions of firm conduct are typically studied in isolation; our paper examines both from the same source of exogenous variation and documents how they interact.

A key focus has been on understanding how firm markups respond to entry or changes in market structure. In the retail sector, the literature has studied Walmart's expansion, finding mixed results: Arcidiacono, Ellickson, Mela, and Singleton (2020) and Ailawadi, Zhang, Krishna, and Kruger (2010) find no significant price response by incumbent grocers, while Atkin, Faber, and Gonzalez-Navarro (2018) document meaningful price declines following Walmart's entry into Mexico. Pricing responses to entry have also been studied in gasoline (Fischer, Martin, and Schmidt-Dengler (2025), Davis, McRae, and Seira (2025), Hosken, McMillan, and Taylor (2008), Chan, Padmanabhan, and Seetharaman (2007a), Chan, Padmanabhan, and Seetharaman (2007b)) or fast food (Pancras, Sriram, and Kumar (2012)). We contribute to this literature by providing causal evidence on how both the intensity and geographic proximity of competition affect retailer markups. We also characterize the nonlinear markup response to increasing numbers of competitors and compare responses across independent retailers and chains.

A related literature studies how competition shapes retailers' product assortment decisions (Watson, 2009; Ren, Hu, Hu, and Hausman, 2011; Matsa, 2011; Oschmann, 2023). For example, Ilanes and Moshary (2019) exploit the deregulation of the liquor markets in Washington and find that retailers expand their assortment in response to an increase in competition. Datta and Sudhir (2013) study how restrictions on spatial differentiation impact product variety, finding that stricter zoning restrictions increase format diversity in the grocery sector. We build on this work by examining how competing retailers differentiate their assortments from one another, distinguishing between horizontal and vertical differentiation in the types of products offered.

A smaller but growing literature studies pricing and assortment responses to competition jointly.

In reduced-form work, Bauner and Wang (2019) show that incumbents react to Walmart and Costco entry by reducing product variety and adjusting their pricing strategies, while Busso and Galiani (2019) find that increased competition in Dominican Republic retail leads to both lower prices and improved service quality. In structural work, Mazzeo (2002) and Singh and Zhu (2008) model the joint determination of product differentiation and equilibrium prices, addressing the endogeneity of market structure through equilibrium entry models.[4] We contribute to this literature by providing causal evidence on how spatial competition jointly shapes pricing and product assortment decisions without imposing assumptions on demand or firm conduct.

Lastly, we add to the growing literature that studies the legalized cannabis industry in the United States (Thomas (2019), Hollenbeck and Uetake (2021), Cirik and Makadok (2023), Hollenbeck and Giroldo (2021), Pavlov (2025)). Both Hollenbeck and Giroldo (2021) and Cirik and Makadok (2023) use the same lottery allocations to, respectively, study the nature of economies of scale for multi-store retailers and the role of first-mover advantage. Our paper is the first to exploit the exogenous variations in local competition created by the lottery. In addition, our simulations provide direct policy implications related to the license cap by quantifying how loosening entry restrictions would reduce prices and shift surplus from firms to consumers.

The rest of the paper proceeds as follows: Section 2 describes the industry background and data with a special emphasis on the lottery allocation process. Section 3 discusses our identification strategy and empirical specifications. Section 4 reports the analyses studying the effect of competition on firm markups and Section 5 reports the results on assortment differentiation. In Section 6, we examine whether and to what extent differentiation is successful in softening price competition. Using the estimation results, Section 7 reports the simulation results on the entry caps. Lastly, Section 8 concludes.

2 Industry and Data

To analyze the effect of local competition on retail markups and product selections, an ideal dataset and setting would have the following key features. First, a source of exogenous variation in the degree of retail competition. Second, high-quality data on retailer prices, wholesale prices, and product availability. Third, it would also contain the full population of competitors and sales to consumers. We study the retail cannabis industry in Washington State, a setting that features each of these. This section describes the institutional details of the license lottery, provides summary statistics on the price and assortment data, and discusses how we compute assortment differentiation between competing retailers.

[4] Other structural analyses of endogenous product choice include Thomadsen (2005), Eizenberg (2014), Draganska, Mazzeo, and Seim (2009), Sweeting (2010), Berry, Eizenberg, and Waldfogel (2016), Fan and Yang (2020), and Wollmann (2018). On the role of vertical relations in shaping product assortments, see Conlon and Mortimer (2021), Hristakeva (2022a), Hristakeva (2022b), and Luo (2024).

2.1 Industry and License Lottery

The retail cannabis industry in Washington State began in 2014 following a state law that authorized the creation of a licensed and taxed cannabis retail sector for adult use. In response to regulatory concerns about excessive retail entry, the state limited the total number of retailers that may enter the market. The total number of retail licenses was capped and allocated across lottery jurisdictions, based on population, population density, and an estimate of past-month marijuana users taken from historical survey data. Lottery jurisdictions typically coincide with city boundaries, though in some cases a jurisdiction may encompass an entire county. The initial cap of 334 licenses across the state was chosen in part to match the number of state-run liquor stores in operation prior to alcohol privatization in 2012. The cap was later increased to 556 in response to excess demand. See Caulkins and Dahlkember (2013) for more on the allocation formula.

Each jurisdiction with excess demand used a lottery to allocate the available retail licenses to applicants. Applicants were assigned a random rank during the lottery, and licenses were awarded to those with the highest ranks, up to the local cap. For example, if a jurisdiction had 3 licenses and received 10 applications, the applicants ranked 1st through 3rd received licenses.[5] Lottery outcomes were announced simultaneously within each jurisdiction. In total, there are 67 jurisdictions that used lotteries to allocate licenses. In jurisdictions with excess demand, the average jurisdiction received 18.8 applications, which competed for approximately 5.9 licenses. Figure 1 shows the distribution of winning probability across lottery jurisdictions, $P_l = \frac{N_l^{cap}}{N_l^A}$, where N_l^A is the total number of applicants in jurisdiction l, and N_l^{cap} is the number of allocated licenses.

A key aspect of the licensing process is that applicants had to specify a proposed store location at the time of application, which we observe in the data. The location requirement, along with reluctance by landlords to permit cannabis shops, made securing a store location a major barrier to entry.[6] The number of licenses available in each jurisdiction, as well as the use of a lottery in oversubscribed markets, was publicly announced online in advance, so applicants knew both how many licenses were available and how excess demand would be allocated. However, applications were submitted simultaneously, and applicants did not know the proposed locations of other firms.

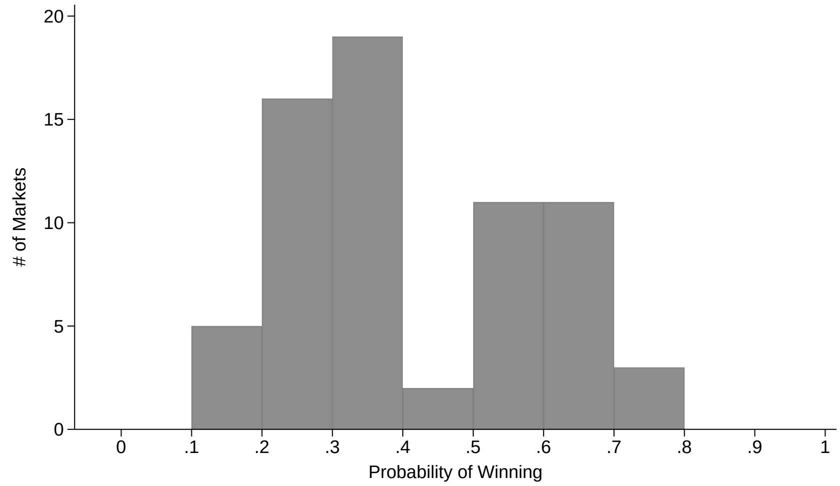
After winning a license, stores were required to enter at the location specified in their application or else go through a cumbersome relocation process. We observe only 3 stores that change locations during the sample period. Compliance with lottery outcomes was high: only 18 lottery winners did not enter the market. Of these, 7 firms passed an initial screening in order to enter the lottery, but

[5]Note that the newly available licenses in 2015 were awarded according to the initial lottery draws; hence, the lottery draws were still the mechanism through which licenses were allocated in the market.

For more details of the legalized cannabis industry in the US, please see, e.g., Hollenbeck and Uetake (2021) and Hansen, Miller, and Weber (2020). This retail lottery is also used by Hollenbeck and Giroldo (2021) to study scale economies in retail and by Cirik and Makadok (2023) to study first-mover effects. In recent years, other states have also used lotteries to allocate licenses, namely Connecticut (2022), Illinois (2023), Maryland (2024), and Missouri (2024).

[6]The regulator imposed strict rules on store locations. For example, retail stores could not be within 1,000 feet of schools, parks, or other public facilities. The application required a proposed store address so that the regulator could verify compliance. For more details, see: `https://lcb.wa.gov/pressreleases/liquor-control-board-approves-lottery-process-retail-cannabis-stores`.

Figure 1: Winning Probability Across Lottery Jurisdictions



Notes: The figure plots the probability of winning a license, defined as (# of Allocated licenses)/(# of Applications), across jurisdictions with a lottery. The unit of observation is at the lottery jurisdiction level.

their applications were rejected after further scrutiny. We observe that stores typically have high variable profits and, while some store owners sell their licenses during the sample period, we do not observe any stores exit. These institutional features support our identifying assumptions, discussed in Section 3.4.

In addition to retailers, the industry also features licensed processors (product manufacturers) and producers (farmers). Unlike retailer licenses, there are no restrictions on the number of processor licenses. As a result, many firms have entered, and the vast majority of producers have also acquired processor licenses. In 2020, we observe 765 licensed processors selling a total of 72.2 thousand unique items across all categories (e.g., usable, extract, edible marijuana). These manufacturers produce a variety of items that retailers may choose from. Note that the state does not allow retailers to own processor licenses. That is, there is no vertical integration between retailers and processors or farmers (City of Tacoma, 2014).

2.2 Data and Sample

We take advantage of a comprehensive administrative database maintained by the Washington State Liquor and Cannabis Board. The database records the universe of transactions in the industry.[7] We observe prices and quantities for each transaction from retailers to consumers. For transactions between manufacturers and retailers, we observe wholesale prices and quantities purchased. There is a gap in the dataset from mid-2017 to mid-2018 when the state switched from one database system to another. Our empirical analyses focus on the data after the switch to maintain data quality. Our sample covers the six quarters from 2018 Q4 to 2020 Q1.

[7]We observe the full population of retailers, and so we observe the full population of legal sales. Direct sales from manufacturers to consumers are barred, as are online sales or retail delivery during the sample period.

Table 1: Summary Statistics

	Mean	SD	5th %ile	95th %ile
Quarterly Retail Profitability				
Retailer's Revenues ($10,000)	59.90	57.33	2.70	176.53
Retailer's Profits ($10,000)	28.85	27.67	1.34	86.34
Prices and Markups				
Retail Prices	16.00	2.66	12.83	19.78
Wholesale Prices	8.13	1.24	6.56	9.69
Markups	7.87	2.18	5.40	10.97

Notes: Summary statistics reflect quarterly information about 397 retailers in 67 lottery jurisdictions over the sample period (2018 Q4-2020 Q1). A product is defined as a producer-by-category pair. Retail and wholesale prices are at the product-by-retailer-by-quarter level and reported in $s.

As explained above, the number of store entrants was regulated by the state, and license caps were allocated separately at the jurisdiction level. Our analyses focus on the 67 jurisdictions, where the license cap was lower than the number of applications even after the expansion of the available licenses in 2015.[8] We observe 397 retailers in these jurisdictions (we use stores and retailers interchangeably). Table 1 summarizes retailers' revenues and variable profits. Retailers are very profitable with average quarterly revenues of about $\$599,000$ and variable profits of $\$288,500$. The largest retailers have roughly 10 times higher sales than the median retailer, earning approximately $4 million per quarter.

One of the key advantages of our data is that we observe wholesale prices that retailers pay to manufacturers; hence, we can calculate markups directly from the data. For our main analyses, we aggregate the data to the product-by-retailer-by-quarter level, where we define a product as a manufacturer-by-product category pair. Prices and markups are then computed by taking quantity-weighted averages over all items sold within each of these product definitions for a given retailer and quarter. The average retail price in our sample is $16.00, while the average wholesale price is $8.13, implying a markup of $7.88. Because regulations bar the use of slotting allowances or other fixed payments to retailers, the observed wholesale prices fully reflect the terms of vertical contracts. While the initial years of the industry, roughly 2014-2016, featured firm entry and large decreases in both wholesale and retail prices, as documented by Hollenbeck and Uetake (2021), by the time period we study, all entries have occurred and upstream and downstream firms are no longer experimenting with prices to learn about demand. Figure 2 plots the average retail and wholesale prices over the sample period, verifying that prices and markups are relatively stable.

Products in the cannabis market fall into several categories. The primary type is usable (flower) marijuana, constituting roughly 51% of sales in 2020. This is an unprocessed plant good that is smoked directly by consumers. This product category is more homogeneous than other categories. The next largest categories are extract products (32.6% of sales), which concentrate active ingredients for consumption via vaporizers, and edible products (7.8% of sales). Table 2 summarizes

[8] Throughout the paper, we use lottery jurisdiction (or jurisdiction) to denote the administrative unit within which a separate lottery was conducted. City refers to an incorporated municipality, which may be a self-contained jurisdiction or be part of a larger county-level jurisdiction.

Figure 2: Average Retail and Wholesale Prices Over The Sample Period

Notes: The figure plots average retail and wholesale prices in $s, averaged across all products and retailers in the sample, over the sample period (2018 Q4-2020 Q1).

Table 2: Summary Statistics by Product Type

Category	*Sale Share %*	*# Producers*	*Price*		*Margin*	
			Avg.	*SD*	*Avg.*	*SD*
Usable Marijuana	51.05	583	15.24	8.59	0.48	0.14
Extract	32.65	485	19.62	7.27	0.47	0.11
Solid Edible	7.85	62	13.91	5.05	0.49	0.12
Infused Mix	3.96	161	8.00	3.63	0.49	0.09
Liquid Edible	2.84	24	17.36	7.68	0.48	0.12
Topical	1.07	58	18.78	7.60	0.48	0.13
Packaged Marijuana Mix	0.58	93	7.93	6.62	0.53	0.12

Notes: The unit of observation is at the product-by-retailer-by-quarter level, where product is defined as a producer-by-category pair.

shares, prices, and markups across all seven product categories. While average retail prices and markups vary across categories, percentage margins (over prices) are relatively uniform, generally around 48%.

Our main analyses look at the causal impact of competition on retailer markups. One may worry that differences in markups stem from differences in wholesale prices rather than competitive pressures. To evaluate the plausibility of this concern, we analyze the variation in wholesale prices across retailers. For these analyses, we look at prices that are aggregated at the retailer-by-quarter-by-item level, where an 'item' is an identifier that is similar to a UPC in the grocery sector. We construct the coefficient of variation (CV) for each item-quarter pair, which captures the variation in wholesale prices across retailers. The median CV is 0 and the mean is just 0.05. Therefore, there is little variation in wholesale prices across retailers. We interpret these patterns as evidence of uniform wholesale prices across retailers, suggesting that differences in retailer markups are not driven by differences in wholesale prices.

Table 3: Summary Statistics of Retail Assortments

	mean	st. dev	5th %ile	95th %ile
Number of Products	93.38	46.97	19.00	170.00
Number of Product Items	1071.54	675.04	122.00	2352.00
Similarity (Product Presence)	0.25	0.11	0.07	0.40
Similarity (Category Share)	0.95	0.08	0.82	1.00
Share of Product Items in Usable	0.37	0.10	0.21	0.50
Diff. in Wholesale Prices	1.12	1.38	0.06	2.83

Notes: The table summarizes retailers' assortment sizes (both in terms of number of products and number of product items supplied) and assortment similarities over quarters. The pairwise similarities capture the retailer's differentiation when compared to its competitor in the closest geographic proximity. Following Hwang, Bronnenberg, and Thomadsen (2010) we "center" the assortment vectors ($A'_r = A_r - \pi_r$) where π_r is the share of products offered by r.

2.3 Measuring Assortment Differentiation

We proxy for retailers' assortment differentiation using both their assortment size and their degree of product overlap with the closest competitor. Table 3 shows that the average retailer sells 93 unique products selected from more than 1,000 product options, suggesting substantial scope for assortment differentiation. We infer that a product is supplied by the retailer in a quarter if the retailer reports sales for the product.

We describe the extent of assortment differentiation across retailers by borrowing from Hwang, Bronnenberg, and Thomadsen (2010). We construct a pairwise assortment similarity measure that captures the degree to which two retailers' assortments overlap in a given time period. For each retailer r, we represent their assortment as a vector $\vec{A_r}$, which contains a dummy for whether or not each product is supplied by the retailer. Omitting time subscripts for ease of readability, the assortment similarity for each pair of retailers r and k is defined as:

$$\text{sim}_{rk} = \frac{\vec{A_r} \cdot \vec{A_k}}{||\vec{A_r}|| \times ||\vec{A_k}||} = \frac{\sum_{j=1}^{N} A_{rj} A_{kj}}{\sqrt{\sum_{j=1}^{N} A_{rj}^2} \times \sqrt{\sum_{j=1}^{N} A_{kj}^2}}, \tag{1}$$

where N is the number of products in the market and $A_{rj} = 1$ if we observe that product j is supplied by retailer r, and zero otherwise. The similarity measure defined in equation (1) is increasing in the share of products sold by both retailers in a period. If both retailers sell an identical set of products, this measure becomes 1, and if they sell non-overlapping assortments, it becomes 0.

As noted by Hwang, Bronnenberg, and Thomadsen (2010), a drawback of this measure is that larger retailers mechanically have greater similarity scores because they carry a larger share of available products. We follow their approach of centering the assortment vectors prior to constructing the similarity. Specifically, for each retailer in each period, we construct $A'_r = A_r - \pi_r$, where π_r is the fraction of all available products carried by retailer r. We then calculate the mean-centered version of the similarity score as:

$$\text{sim}'_{rk} = \frac{\vec{A'_r} \cdot \vec{A'_k}}{||\vec{A'_r}|| \times ||\vec{A'_k}||}. \tag{2}$$

Table 3 summarizes the pairwise similarity in assortments defined by equation (2). For each retailer, we summarize its assortment similarity to the competitor in closest geographic proximity. We find that product-level similarity is markedly lower than in the grocery sector, where Hwang, Bronnenberg, and Thomadsen (2010) and Hristakeva (2022b) both report values above 0.7. The comparison suggests that retailers in the cannabis market select products strategically to differentiate themselves from local competitors.

To capture differentiation at the category level, the product-level approach is uninformative because almost all retailers supply at least one product in each category. Instead, we construct a category-level similarity measure by applying equation (2) to vectors of category shares rather than product indicators. That is, each entry of the vector records the share of a retailer's items in a given category (e.g., 0.2 if 20% of retailer r's items are in category j). The average category-level similarity is 0.95, reflecting the fact that most retailers allocate their assortments across categories in broadly similar proportions.

We construct two additional variables used in Section 5 to evaluate whether assortment differentiation reflects vertical or horizontal positioning. The first variable captures the share of a retailer's assortment accounted for by the usable marijuana category. As usable items are relatively homogeneous and offer limited brand differentiation, a lower share of usable products supplied is interpreted as evidence of greater horizontal differentiation. Table 3 shows that, on average, 37% of a retailer's assortment falls within this category. The second variable, summarized in the last row of Table 3, proxies for vertical differentiation by comparing average wholesale prices. For each retailer, we compute the simple average of wholesale prices in its assortment, and then take the absolute difference in this average relative to its closest competitor. This captures whether neighboring retailers offer products at systematically different wholesale price levels.

3 Empirical Strategy

The main challenge in estimating the causal effect of competition on markups (and product differentiation) is that market structure is endogenous. Firms do not enter markets randomly; rather, they base their decisions on expected demand and cost conditions. This creates a correlation between the number of competitors and unobserved local market characteristics, complicating the interpretation of any observed relationship between competition and firm behavior. In the case of markups, the key concern is that higher-demand areas are likely to support both higher prices and a larger number of firms.

We address this endogeneity by exploiting the license lotteries described in Section 2, following the approach of Borusyak and Hull (2023). This section presents our identification strategy, describes the competition measures we construct, provides formal tests of the lottery's randomization, and summarizes the key assumptions.

3.1 Identification Strategy

We first present a simplified representation of our main specification to explain the identification problems. Let Markup_r denote retailer r's average markup and Comp_r track a measure of local competition faced by retailer r, such as the number of nearby competitors within a given radius or the distance to the closest competitor. The object of interest is β, which captures the causal effect of competition on markups:

$$\text{Markup}_r = \beta\,\text{Comp}_r + \varepsilon_r, \tag{3}$$

where ε_r captures unobserved determinants of markups, including local demand and cost conditions. Naïve OLS estimation of β is biased because areas with favorable unobserved conditions attract both more firms and higher markups, inducing a correlation between Comp_r and ε_r.

The advantage of our empirical setting is that the lottery assignment provides exogenous variation in the number and locations of competitors in local areas. However, the lottery does not fully randomize competition. While license allocation is randomized among applicants, the locations of applicants are not. Firms selected their proposed locations based on expected profitability before the lottery occurred, meaning that areas with more applicants may have been perceived as more attractive. We recognize that non-random exposure to the exogenous lottery shock leads to systematic differences in treatment assignment across observations, introducing omitted variable bias.

The key insight for identification is that realized competition is determined by a known function $f(g, w)$, where w denotes the configuration of applicants (their number and locations) and g denotes the vector of lottery outcomes. That is, the realized competitive intensity, Comp_r, is a *composite variable*, determined by:

1. Non-random determinants: The endogenous applicant pool (w), which reflects firms' strategic location choices.

2. A true experiment: The lottery outcomes (g), which provide random variation in which applicants receive licenses; and

While the lottery randomizes license allocation, it does not fully randomize the level of competition because applicant location w reflects endogenous entry decisions.

Following Borusyak and Hull (2023), we decompose realized competition as:

$$\text{Comp}_r = \mu_r + \underbrace{(\text{Comp}_r - \mu_r)}_{Z_r}, \tag{4}$$

where $\mu_r \equiv E[\text{Comp}_r \mid w]$ is the expected competition that retailer r would face on average across all possible lottery draws, conditional on the set of nearby applicants. It captures the component of realized competition that is predictable from the pre-lottery information on the number of applicants and their locations. For example, when competition is measured as the number of competitors within a given radius, $\mu_r = P_l \cdot N_r^A$, where P_l is the winning probability in the lottery jurisdiction where r

is located and N_r^A is the number of applicants in the neighborhood around r.

The *recentered shock* $Z_r \equiv \text{Comp}_r - \mu_r$ captures whether retailer r ended up with more or fewer competitors than expected. By construction, $E[Z_r \mid w] = 0$, and this is precisely the variation arising purely from lottery randomness, which we use to estimate the causal effect of competition on firm strategies.

Specifically, our preferred specification regresses markups on realized competition while controlling for expected competition:

$$\text{Markup}_{jrmt} = \beta \text{Comp}_r + \delta \mu_r + \theta_{jt} + \eta_{jm} + \varepsilon_{jrmt}, \tag{5}$$

where Markup_{jrmt} is the retail markup (in dollars) for product j at retailer r in city m and quarter t. We include quarter-by-product fixed effects, θ_{jt}, to control for time-varying fluctuations in product demand and cost that affect all retailers, while the city-by-product dummies, η_{jm}, capture city-specific product demand conditions, ensuring that broader market-level variation does not confound our estimates. Including μ_r as a control absorbs the systematic component of competition that varies with endogenous applicant locations, so that β is identified from lottery-driven variation alone. Standard errors are clustered at the city-by-quarter level throughout.

We present Equation (5) as our baseline specification for the markup analyses in Section 4. Sections 5 and 6 adapt this framework to study assortment decisions and product-level competition, respectively.

Lottery Identification Identification of β in Equation (5) relies on the exogeneity of the lottery shock. Our key identifying assumption (A1) is:

$$E[\varepsilon_r \mid g, w] = E[\varepsilon_r \mid w], \tag{6}$$

where we omit the $\{jmt\}$ subscripts without loss of generality. Equation (5) states that, conditional on who applied and where, the lottery outcome is uninformative about markup shocks in expectation. This assumption satisfied by the institutional design of the lottery, which randomly ranked applicants and awarded licenses to those above a predetermined cap.

Under Equation (6), the recentered shock is orthogonal to unobserved markup determinants:

$$E[Z_r \, \varepsilon_r] = 0. \tag{7}$$

This follows by iterated expectations.[9]

To make the roles of each variable transparent, note that substituting $\text{Comp}_r = \mu_r + Z_r$ into

[9] Since Z_r is a function of (g, w) and $E[Z_r \mid w] = 0$ by construction: $E[Z_r \varepsilon_r] = E\big[E[Z_r \varepsilon_r \mid g, w]\big] = E\big[Z_r E[\varepsilon_r \mid g, w]\big] = E\big[Z_r E[\varepsilon_r \mid w]\big] = E\big[E[Z_r \mid w] \cdot E[\varepsilon_r \mid w]\big] = 0$. The second equality uses measurability of Z_r with respect to (g, w); the third applies Equation ((6)); the fourth applies the law of iterated expectations conditioning on w. See Borusyak and Hull (2023), Appendix B.1.

Equation (5) yields:

$$\text{Markup}_{jrmt} = \beta Z_r + (\beta + \delta)\mu_r + \theta_{jt} + \eta_{jm} + \varepsilon_{jrmt},$$

This reparameterization shows that β is identified from the exogenous lottery shock Z_r, while μ_r absorbs the relationship between expected competition and markups that may reflect confounding.

Instrumental Variable Approach We note that the same shock exogeneity assumption (A1) also supports an instrumental variables approach, using the recentered shock Z_r as an instrument for Comp_r. The instrument is relevant because $E[Z_r \, \text{Comp}_r] \neq 0$, and exogenous by Equation (7). First-stage results for the IV specification are reported in Appendix A.3. For completeness, Section 4 estimates the IV analogue of our main specification, using:

$$\text{Markup}_{jrmt} = \beta \text{Comp}_r + \theta_{jt} + \eta_{jm} + \varepsilon_{jrmt}, \tag{8}$$

and instrumenting Comp_r with Z_r.

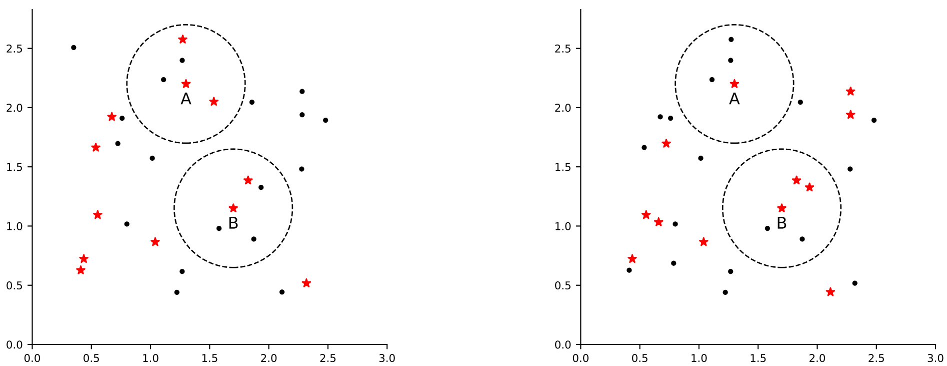
Stylized Example To build intuition for the identification strategy, consider a lottery jurisdiction that received 30 applications for retail licenses and ultimately granted 12, implying a winning probability of 0.4 for each applicant. Panel A of Figure 3 illustrates the spatial distribution of applicants, marked by black dots, and the locations of lottery winners, indicated by red stars. Consider two license winners, Retailer A and Retailer B, around which we draw circles with 0.5km radii. Both are located in areas that received 5 applications, meaning expected competition is the same: $\mu_A = \mu_B = 5 \times 0.4 = 2$. However, due to the randomness of the lottery, Retailer A ends up with two competitors within the radius ($Z_A = 0$), while Retailer B has only one ($Z_B = -1$). An alternative lottery draw could have led to the market structure shown in Panel B of Figure 3, where Retailer A would have had no competitors ($Z_A = -2$) and Retailer B would have had two ($Z_B = 0$). Our identification strategy controls for expected competition μ_r, isolating the variation in markups driven exclusively by the lottery.

3.2 Measures of Competition and Implementation

We measure the degree of local competition using three variables: (i) the distance to the nearest competitor's store, (ii) the number of competitors within a 0.5km radius, and (iii) the number of competitors within a 1km radius. To construct these measures, we geocode all store addresses and compute pairwise distances using the geographic coordinates of each location. Based on this distance matrix, we compute, for each store, the distance to its closest competitor, as well as the number of other retailers within 0.5km and 1km radii.

Table 4 summarizes these measures. These variables do not vary over time because all stores entered prior to the start of our sample period. The distance to the closest competitor shows

Figure 3: Observed and Hypothetical Market Structures



(a) Observed Lottery Winners (b) Lottery Winners with Alternative Draw

Notes: Black dots indicate applicant locations; red stars indicate lottery winners. Circles show 0.5km radii around Retailers A and B.

Table 4: Measures of Local Competition and Expected Competitive Pressures

	mean	st. dev	5th %ile	95th %ile
Competitors				
Dist. to Closest Comp.	1.72	2.58	0	6
# Comp. in 0.5km	0.69	0.94	0	3
# Comp. in 1km	1.08	1.38	0	4
E[Competitors]				
E[Dist. to Closest Comp.]	2.16	2.78	0	7
E[# Comp. in 0.5km]	0.62	1.04	0	3
E[# Comp. in 1km]	0.94	1.27	0	4

Notes: The unit of observation is at the retailer level. For each retailer, we summarize the observed competitive environment (e.g. # of competitors in a 1km radius) and the expected level of competition μ_r in the local area. For count measures, μ_r = # of applicants in radius $\times P_l$. For distance, μ_r is the average distance to the closest competitor across 500 simulated lottery draws.

substantial variation in spatial competition: the average is 1.72km with a standard deviation of 2.58. Retailers face, on average, fewer than one competitor within 0.5km (0.69) and about one competitor within 1km, with substantial variation across stores. We considered other radii for measuring the number of local competitors, and our choice of 0.5km and 1km is motivated by results on spatial competition described in Section 4.

As outlined above, our identification strategy requires constructing expected competition, μ_r, for each competition measure. When analyzing the effect of the number of competitors within a given radius, we calculate the expected number of competitors as the number of applicants within that radius multiplied by the winning probability in its lottery jurisdiction. For example:

$$E[\#\text{ of comp. in }0.5\text{km}_r] = (\#\text{ of applicants in }0.5\text{km}_r) \times P_l, \tag{9}$$

where P_l is the winning probability in the jurisdiction where retailer r is located.

For the distance to the nearest competitor, we simulate μ_r as:

$$E[\text{Distance to closest comp.}_r] = \frac{1}{500}\sum_{sim=1}^{500} \text{dist closest}_r^{sim}, \tag{10}$$

where each simulation draw sim assigns licenses randomly according to the winning probabilities of the relevant lottery jurisdiction and computes the distance from retailer r to its nearest competitor.

The second panel of Table 4 summarizes the variables tracking expected competitive pressures. The realized and expected values match closely: for example, E[# Comp. in 1km] is 0.94 and the realized value is 1.08. One potential concern is that our drawn neighborhoods may cross the boundaries of lottery jurisdictions with different winning probabilities. For instance, if a high-demand market is adjacent to a low-demand one, applicants might strategically locate near the border within the low-demand market. Our use of narrow radii (0.5km and 1km) reduces the scope for this concern.

Table 5: Variation in Recentered Shock

	App (jurisdiction)	Win prob	μ_r	mean Z_r	SD Z_r
# App (1km)					
1	36.15	0.43	0.43	0.41	1.01
2	28.63	0.42	0.85	0.33	1.22
3	59.50	0.38	1.14	0.31	0.96
4	34.44	0.33	1.32	-0.07	1.27
5	74.86	0.29	1.43	0.48	1.55
6	46.89	0.27	1.61	0.84	1.65
7	57.60	0.32	2.22	-0.62	1.30
8	130.76	0.27	3.43	-0.90	1.60

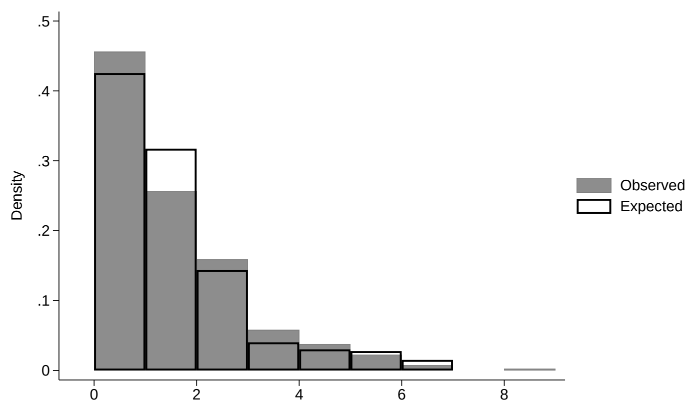
Notes: The unit of observation is at the retailer level. Retailers are grouped by the number of applicants located within 1km of the focal store. For each bin, we report the average number of applicants in the lottery jurisdiction, the average win probability, the expected number of competitors ($\mu_r = N_r^A \times P_l$), and the mean and standard deviation of the recentered shock $Z_r = \text{Comp}_r - \mu_r$.

To characterize the identifying variation in our data, Table 5 summarizes the distribution of the recentered shock Z_r across retailers grouped by the number of nearby applicants within 1km. Lottery jurisdictions with more nearby applicants have, on average, lower win probabilities and higher expected competition (μ_r). The dispersion of Z_r is broadly similar across bins, indicating that identifying variation is not concentrated in a narrow set of high-density locations, and the mean of Z_r remains close to zero relative to its standard deviation.

3.3 Formal Tests of Randomization

Our main identifying assumption requires that the lottery produced random license allocations. We formally test whether the lottery indeed generated random numbers of competitors conditional on the number of applicants using a χ^2 test. The test assesses whether the realized distribution of local competition and the expected distribution implied by the lottery assignment mechanism are statistically indistinguishable. We compare the realized number of competitors within a 1km radius to the expected number of competitors. Figure 4 shows the distributions of the expected and realized number of competitors within a 1km radius. To match the discrete nature of the variable tracking the observed number of competitors, we round the value of the expected number of competitors to the nearest integer. The test produces a p-value of 0.481, failing to reject the null hypothesis that the observed number of competitors is consistent with random assignment. We interpret these results as reassurance that the distribution of the number of local competitors is truly random once we condition on the number of nearby applicants.

Figure 4: Distribution of # Competitors in 1km Radius



Notes: The unit of observation is at the retailer level. The figure tabulates the number of competitors observed within 1km of each focal store and the expected number of competitors within 1km of each store. The expected number of competitors in a 1km radius is constructed as # of applicants in 1km $\times$ winning probability in the lottery jurisdiction, rounded to the nearest integer. The χ^2 test of these distributions produces a p-value of 0.481.

We also test for balance in observable market characteristics. If the lottery shock is truly exogenous, realized competition should be uncorrelated with local demographics after conditioning on expected competition. For each store, we observe the demographic characteristics of its census tract and estimate:

$$\text{Demo}_r = \alpha + \gamma \text{Comp}_r + \psi \mu_r + \nu_r, \tag{11}$$

where Demo_r denotes a census tract characteristic (population, average age, and income). Table 6 reports these balance tests. We find no significant relationship between local market characteristics and either the distance to the closest competitor (columns (1)-(3)) or the number of local competitors (columns (4)-(6)). This suggests that the selection concern may be limited.

A separate concern is whether different types of locations attracted systematically different types of applicants. If more sophisticated or risk-tolerant firms targeted dense areas, post-entry differences in pricing or differentiation could reflect firm type rather than competitive pressure. Appendix A.1 examines whether there are significant and systematic relationships between competition and observable characteristics of applicants and finds none. In addition, we show in Appendix A.2 results

Table 6: Randomization across Demographics

	Population (1)	Age (2)	Income (3)	Population (4)	Age (5)	Income (6)
Dist. to Closest Comp.	28.15	0.14	991.45			
	(32.00)	(0.15)	(709.54)			
E[Dist. to Closest Comp.]	–30.92	–0.01	–1,020.94			
	(29.72)	(0.14)	(659.01)			
# Comp (1km)				13.11	–0.26	–900.35
				(45.62)	(0.21)	(1,011.84)
E[# Comp (1km)]				–44.62	0.30	–374.48
				(51.17)	(0.24)	(1,134.87)
N	397	397	397	397	397	397

Notes: The unit of observation is at the retailer level. The dependent variable is the demographic profile of the census tract where the retailer is located measured by population, average age, and income. Columns (1)-(3) use distance to the closest competitor as the competition measure; columns (4)-(6) use the number of competitors within 1km. All regressions control for the expected value of the respective competition measure.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

from an alternative strategy that isolates variation in realized competition within more homogeneous competitive environments by estimating the effect of realized competition on markups separately for subsamples defined by the expected number of nearby competitors. We find consistently negative effects for each subsample, with no clear pattern in the relationship between effect size and the expected level of competition.

3.4 Identifying Assumptions

We summarize the maintained assumptions below:

A1 Random license assignment (Shock Exogeneity, Assumption 1 in Borusyak and Hull (2023)). Conditional on the (pre-lottery) applicant pool w, the realized lottery outcome g is mean-independent of unobserved markup determinants (Equation 6). This is satisfied by the institutional design of the Washington State lottery, which ranked applicants randomly and awarded licenses to those above a predetermined cap. Section 3.3 provides formal tests.

A2 Known assignment process (Assumption 2 in Borusyak and Hull (2023)). We know the proposed locations of each license application; therefore, expected competition μ_r is correctly specified as the conditional expectation of Comp_r given w.

A3 No post-lottery relocation or exit. Retailers enter at their proposed locations. This is institutionally supported as the state required applicants to enter at the address specified in their application, and we observe only 3 relocations among 397 lottery winners. Non-compliance was minimal: only 18 lottery winners did not enter, of whom 7 were rejected on further regulatory scrutiny. No stores exit during the sample period.

A4 Partial interference (SUTVA). A retailer's markup depends on competitors within the specified radius (0.5km or 1km) but is not affected by lottery outcomes of more distant applicants. We provide evidence supporting this assumption in Appendix A.7, which shows that "indirect" competitors located 1-3km away have small and statistically insignificant effects on markups.

Note that the identification strategy is valid whenever the above assumptions hold, regardless of winning probabilities or applicant density. The precision of our estimates, however, depends on the variance of the lottery-induced shock. For example, when competition is measured as the number of nearby competitors, $\text{Var}(Z_r) = N_r^A \cdot P_l(1 - P_l)$, which is increasing in the number of nearby applicants and maximized when $P_l = 0.5$.[10] The first-stage results in Appendix A.3 confirm that the lottery generates meaningful variation in local competition.

In addition, as in any pooled regression, the estimator implicitly weights observations by their contribution to identifying variation. If the effect of competition varies across markets, the pooled coefficient is a variance-weighted average of local effects rather than a simple average. We investigate this possibility and find no evidence of economically meaningful treatment effect heterogeneity. Estimated effects are stable across market types, and reweighting observations to equalize each market's contribution produces estimates that are quantitatively similar to the baseline. In Appendix A.10, we show the reweighted estimates and show visually that identifying variation is not disproportionately concentrated or systematically related to market features.

4 Retailers' Price Response to Local Competition

In this section, we present causal estimates of how spatial competition impacts retail markups. Textbook economic models suggest that increased competition results in lower prices. The magnitude and shape of this decline, however, depend on the nature of competition and underlying demand conditions. For example, under Cournot competition, each additional entrant imposes a smaller marginal impact on price, leading to a concave relationship between competition and markups. Alternatively, Bertrand competition with differentiated products suggests sharper responses to nearby rivals, especially when firms are closely substitutable. These differences in theoretical predictions underscore the need for empirical estimates that do not rely on assumptions about demand or firm conduct. Exploiting exogenous variation in local competition generated by the license lottery, we show how and to what extent markups respond to local competitive pressure. Specifically, we evaluate the relationship between markups and competitive pressures using distance to the competitor in closest geographic proximity and the number of competitors within 0.5km and 1km radii (summarized in Table 4).

Distance to Closest Competitor Table 7 shows our regression results using distance to the competitor in closest geographic proximity as our measure of local competition. Column (1) reports the naïve OLS estimate without controlling for expected distance. We see that retailers charge higher

[10] The exact variance is $\text{Var}(Z_r) = N_r^A \cdot P_l(1 - P_l) \cdot \frac{N^A - N_r^A}{N^A - 1}$, reflecting that lottery draws are without replacement.

Table 7: Effect of Retail Competition on Markups: Distance to Closest Competitor

	(1) OLS	(2) Control	(3) IV
Dist. to Closest Comp.	0.131***	0.309***	0.455***
	(0.029)	(0.069)	(0.114)
E[Dist. to Closest Comp.]		–0.237***	
		(0.068)	
N	194,509	194,509	194,509
City-product FE	yes	yes	yes
Quarter-product FE	yes	yes	yes

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retail markups measured in dollars. Regressions include FE at the level of the quarter-by-product and city-by-product. Column (1) reports OLS without controlling for expected competition. Column (2) adds the expected distance to the closest competitor as a control. Column (3) instruments realized distance with the recentered shock Z_r=Distance to closest comp. - E[Distance to closest comp.]. The Kleibergen-Paap F statistic is 123.53. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

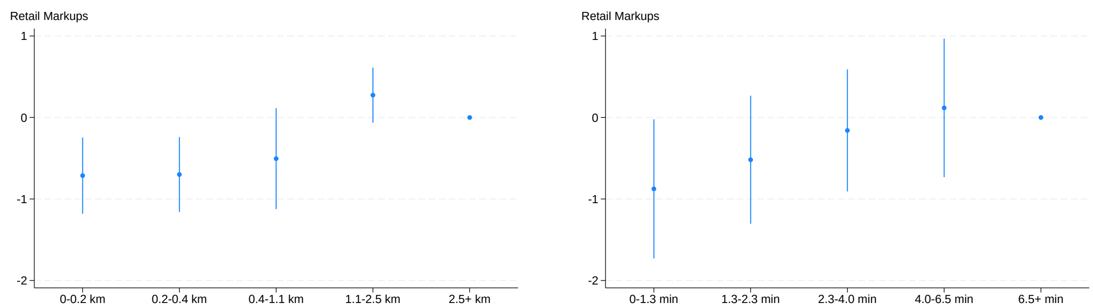
markups when their nearest competitor is farther away. However, this estimate is likely attenuated due to the endogeneity concerns discussed in the previous section. Column (2) shows our preferred specification from Equation 5, where we control for expected distance to the closest competitor. After adding these controls, the parameter estimate on distance to the closest competitor more than doubles from 0.131 to 0.309, confirming that unobserved demand conditions in Column (1) create a positive correlation between competitor proximity and markups. Our preferred estimate implies that moving a store's closest competitor 1km farther away increases average markups by \$0.309, or about 4% of the mean markup. Column (3) confirms this result using the IV strategy from Equation 8, providing additional evidence that proximity to competitors significantly constrains retailer pricing power. As a robustness check, we verify in Appendix A.6 that the estimates are similar when excluding Seattle, as the largest market may have different competitive dynamics.

While this gives an average effect, our next set of results allows us to delineate the extent to which price competition varies with distance to the closest competitor. To explore nonlinearities in the relationship, we replace the linear competition term in Equation 5 with group indicators:

$$\text{Markup}_{jrmt} = \sum_{a=1}^{G} \beta_a \mathbb{I}\{\text{Comp}_r = a\} + \sum_{a=1}^{G} \delta_a \mathbb{I}\{\mu_r = a\} + \eta_{jm} + \theta_{jt} + \varepsilon_{jrmt}, \tag{12}$$

estimating a separate coefficient β_a for each level of realized competition. For our distance measure, we discretize the distance variable into five groups with an equal number of retailers in each group: retailers whose closest competitor is within 0-0.2km, 0.2-0.4km, 0.4-1.1km, 1.1-2.5km, and >2.5km. The first panel of Figure 5 shows the estimated effects on markups for each distance band. The excluded group is comprised of retailers whose nearest competitor is at least 2.5km away. We find a decreasing and somewhat nonlinear relationship: the presence of a competitor within 0.2km or 0.4km has a very large impact on markups, and as the closest competitor becomes further away

Figure 5: Effects of Competition on Markup



(a) Dist. to the Closest Comp.

(b) Travel Dist. to the Closest Comp.

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Panel (a) uses geographic distance (km); panel (b) uses travel duration (minutes). Results include FE for quarter-by-product and city-by-product. We control for expected distance to the closest competitor using dummies for the discretized expected distance. The excluded category is retailers whose closest competitor is more than 2.5km away (panel a) or more than 6.5 minutes away (panel b). Standard errors are clustered at the city-by-quarter level. Appendix A.5 reports a table with the estimated coefficients.

from the focal retailer, the effect fades out. Specifically, retailers whose closest competitor is further than 2.5km away charge approximately \$0.70 higher markups than retailers facing a competitor within a 0.2km distance, a difference that corresponds to about 9% lower markups from the mean or 0.3 standard deviations. The graph also shows that having the closest competitor within 1.1-2.5km or farther than 2.5 km has the same estimated effect on markups. These results highlight that retail competition in this market is localized as retailers' markups respond strongly when the closest competitor is within approximately 1km distance.

As an alternative specification, we use travel time in place of physical distance. Panel (b) in Figure 5 shows that our main results do not depend on the way we capture proximity to the closest geographically located competitor. When the closest competitor is within 1.3 minutes of driving time, retailers charge markups that are approximately \$0.87 lower than if it is farther than 6.5 minutes away. Similarly, the effect of competition on markups diminishes quickly with travel duration: on average, the effect of a competitor on markups is the same irrespective if the closest competitor is located 4-6.5 minutes or farther than 6.5 minutes away. Appendix A.4 reports additional results using travel duration. Given the similarity between the two measures of distance, the rest of the analyses rely on our preferred specification using distance to the closest competitor.

Number of Competitors In addition to geographic proximity to a competitor, the local competitive pressures may change as the number of competitors in an area increases, a relationship that we analyze below. Building on our results that competition in this market is highly localized, we estimate the effects of the number of competitors within a radius of 0.5 and 1 kilometer, respectively. Table 8 presents the results. We use 0.5 kilometer radius for columns (1)-(3) and 1 kilometer for columns (4)-(6).

Two patterns emerge. First, the results again demonstrate the role of endogeneity: unobserved

Table 8: Effect of Retail Competition on Markups: Number of Competitors

	(1) OLS	(2) Control	(3) IV	(4) OLS	(5) Control	(6) IV
# Comp (0.5km)	–0.411***	–0.454***	–0.612***			
	(0.048)	(0.053)	(0.120)			
# Comp (1km)				–0.261***	–0.298***	–0.358***
				(0.023)	(0.035)	(0.067)
E[# Comp] (0.5km)		0.079*				
		(0.047)				
E[# Comp] (1km)					0.062*	
					(0.038)	
N	194,509	194,509	194,509	194,509	194,509	194,509
City-product FE	yes	yes	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes	yes	yes

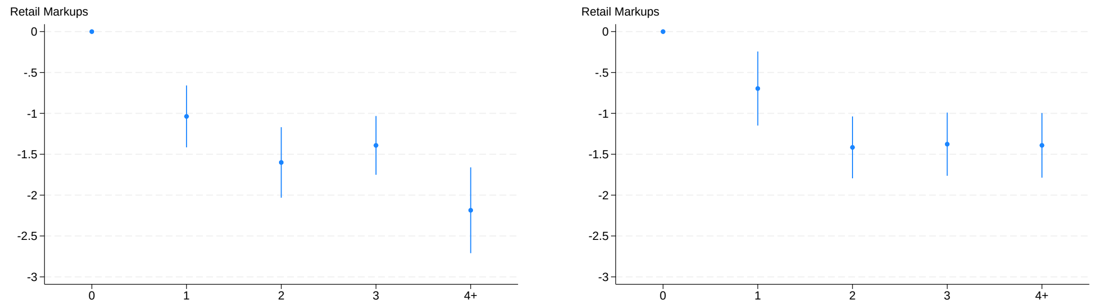
Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Columns (1)-(3) measure competition within a 0.5km radius; columns (4)-(6) within 1km. Results include FE for quarter-by-product and city-by-product. The instrumental variables in columns 3 and 6 are Z=# Comp. in radius - E[# Comp. in radius]. The Kleibergen-Paap F statistics for each competitive variable are 46.78 and 117.54, respectively. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

local demand or cost conditions attenuate the naïve OLS estimates toward zero. For example, the parameter estimate on the number of competitors within 0.5km decreases from -0.411 to -0.454 between columns (1) and (2). Second, the effect of an additional competitor attenuates as we expand the radius from 0.5 to 1km, with the parameter estimates changing correspondingly from -0.454 to -0.298.

Following Bresnahan and Reiss (1991), we next ask how much of the competitive effect is driven by the presence of the first competitor versus subsequent ones. That is, how much larger is the impact of moving from a local monopoly to a duopoly than moving from 2 to 3 firms, etc.? Figure 6 re-estimates the controlled OLS specification from columns (2) and (5) using Equation 12, with a separate coefficient for each value of the number of competitors. The excluded category consists of retailers without any competitors within 1.5km, so that the estimates are comparable across the graphs.

In both cases, we find that the effects of the first and second competitors are similar in magnitude, whereas increasing the number of nearby competitors beyond the second one has close to zero additional impact on markups. Our finding that most of the changes in competitive conduct occur with the entry of the first and second competitors aligns with the findings of Bresnahan and Reiss (1991). In our setting, we are able to confirm these relationships directly using precise information on retailers' markup and leveraging stricter market definitions based on geographic proximity.

Figure 6: Effects of Number of Competitors on Markup



(a) Number of Comp. within 0.5km

(b) Number of Comp. within 1km

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Panel (a) uses competitors within 0.5km; panel (b) within 1km. Results include FE for quarter-by-product and city-by-product. We control for expected competition using dummies for the (rounded) expected number of competitors within the relevant radius. The excluded category is defined as retailers without any competitors within 1.5 kilometers. Standard errors are clustered at the city-by-quarter level. Appendix A.5 reports a table with the estimated coefficients.

4.1 Heterogeneity by Store Type: Chains vs. Independents

We next explore whether the competitive effects documented above vary across retailer types by comparing independent stores to those that are part of a multistore chain. In our sample, 223 of the stores are affiliated with a chain, and 174 are independent. In this market, chains arise when applicants submit multiple applications and win more than one license. As of 2025, chains are limited to owning no more than 5 stores. These strict ownership rules, combined with a limited resale of licenses, imply that nearly all multi-store ownership reflects the initial lottery allocation.[11]

To evaluate whether chain affiliation moderates pricing responses to competition, we replicate our main markup regressions separately for chain and independent retailers. As with our main competition measures, chain status is determined by both lottery draws (random) and the owner's decision to submit multiple applications (endogenous). We construct an indicator for whether the expected number of stores an owner wins exceeds one, $\mathbb{I}\{E[\#\text{Stores}] > 1\}$. The empirical specification follows Equation 5 with the inclusion of $\mathbb{I}\{E[\#\text{Stores}] > 1\}$ as an additional control following the same recentering logic. Results are shown in Table 9.

Comparing the estimates across the two groups reveals a clear difference: Independent stores consistently exhibit a stronger response to competition, with markups falling more steeply as the number or proximity of nearby competitors increases. Chain stores also reduce markups in response to competition, but the magnitude of the response is smaller. These patterns hold even though most chains in our data include only 2 or 3 stores, suggesting that even modest multi-store ownership attenuates the sensitivity of pricing to local competitive conditions.

[11] Retailers might react to increased competition by choosing to form chains to increase differentiation with greater branding (Bronnenberg, Dubé, and Moorthy (2019), Hollenbeck (2017).) For the full joint distribution of number of applications filed and number of post-lottery stores owned, see Hollenbeck and Giroldo (2021).

Table 9: Effect of Retail Competition on Markups: By Store Type

	Independent Stores			Chain Stores		
	(1)	(2)	(3)	(4)	(5)	(6)
Dist. to Closest Comp.	0.210**			0.160**		
	(0.105)			(0.071)		
# Comp (0.5km)		–0.922***			–0.302***	
		(0.129)			(0.061)	
# Comp (1km)			–0.516***			–0.168***
			(0.054)			(0.032)
E[Dist. to Closest Comp.]	0.019			–0.058		
	(0.111)			(0.040)		
E[# Comp] (0.5km)		–0.127**			0.097***	
		(0.059)			(0.030)	
E[# Comp] (1km)			0.070			0.028
			(0.056)			(0.020)
N	64,933	64,933	64,933	103,605	103,605	103,605
City-product FE	yes	yes	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes	yes	yes
$\mathbb{I}\{E[\#\text{Stores}] > 1\}$	yes	yes	yes	yes	yes	yes

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Columns (1)-(3) report results for independent stores; columns (4)-(6) for chain stores. Results include FE for quarter-by-product and city-by-product. The regressions also include a dummy equal to 1 if the expected number of stores for the license retailer is greater than 1. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These findings align with two potential mechanisms. First, prior work in this setting shows that chains differentiate themselves by offering larger assortments, which may reduce the intensity of direct price competition (Hollenbeck and Giroldo (2021)). Second, chains may employ uniform pricing policies that decouple prices from local demand and competitive conditions. This practice has been documented across a variety of retail contexts (DellaVigna and Gentzkow (2019), Hitsch, Hortaçsu, and Lin (2021)). While we do not attempt to disentangle these effects, our findings are consistent with both playing a role.

These results highlight the importance of retailer heterogeneity in shaping competitive dynamics. Markets composed primarily of independent retailers may see more intense price competition in response to entry, whereas areas dominated by chain stores may see more muted effects. This has implications for policy debates in this industry, with some states restricting the retail sector to consist of small "mom and pop" operations and other states allowing for very large retail chains.[12] These results imply that the restrictions on maximum chain size imposed by the regulator may have incidentally increased local price competition, partially offsetting the competitive limits imposed by the license cap.

[12] For example, New York and Massachusetts allow no more than 3 retail licenses per licensee, while Colorado and California do not explicitly set a cap.

Table 10: Effect of Competition on Retail Differentiation

	Assort Size (1)	Product Assort Diff (2)	Category Assort Diff (3)	Usable Share of Assort (4)	abs(Wholesale Price Diff) (5)
Dist. to Closest Comp.	–7.994***	0.007***	0.003**	0.006***	0.018
	(0.799)	(0.003)	(0.001)	(0.002)	(0.018)
E[Dist. to Closest Comp.]	7.656***	0.020***	0.000	–0.000	–0.090**
	(1.043)	(0.004)	(0.002)	(0.002)	(0.037)
Assort. Size (log)		0.048***	0.039***	0.026***	–0.304***
		(0.004)	(0.005)	(0.010)	(0.085)
N	1,935	1,935	1,935	1,935	1,929
City FE	yes	yes	yes	yes	yes
Quarter FE	yes	yes	yes	yes	yes

Notes: The unit of observation is at the retailer-by-quarter level. In columns (2)-(5), dependent variables are defined relative to the retailer's nearest competitor (see text). Columns (2)-(5) control for assortment size. Results include FE for city and quarter. Standard errors are clustered at the city-by-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Competition and Assortment Differentiation

Beyond pricing, retailers may respond to competitive pressure by adjusting their product assortments. Standard models of horizontal differentiation predict that firms facing closer rivals will differentiate to soften price competition (d'Aspremont, Gabszewicz, and Thisse (1979); Ginsburgh, de Palma, Papageorgiou, and Thisse (1985)), though the direction and magnitude of assortment responses are ultimately empirical questions. Our focus on assortment strategy is motivated by evidence in marketing that store assortments are central to retailer positioning (Ailawadi and Keller (2004)) and are as important as prices in influencing store choice (Fox, Montgomery, and Lodish (2004), Briesch, Chintagunta, and Fox (2009), Broniarczyk, Hoyer, and McAlister (1998), Hoch, Bradlow, and Wansink (1999), Bonfrer, Chintagunta, and Dhar (2022)).[13]

For these analyses, our dependent variables track assortment decisions, which are constructed at the retailer-by-quarter level. The identification approach mirrors that used in the markup analysis, relying on exogenous variation in realized competition once we control for expected competitive pressures. Specifically, we estimate the following regression:

$$\text{Assort}_{rmt} = \beta \text{Comp}_r + \delta \mu_r + \alpha_t + \gamma_m + \varepsilon_{rmt}, \tag{13}$$

where Assort_{rmt} denotes the assortment outcome for retailer r in city m at time t. Comp_r is realized competition, and μ_r is expected competition. The specification includes quarter fixed effects, α_t, and city fixed effects, γ_m. Standard errors are clustered at the city-by-quarter level.

[13] Another potential margin of differentiation is advertising. In our data, however, advertising variation is limited: the majority of spending comes from Seattle, and two-thirds is allocated to outdoor billboards. We do not examine advertising in this paper.

Our first set of results uses assortment size as the dependent variable, which we operationalize as the number of unique products supplied by the retailer in that quarter. Column (1) in Table 10 shows that retailers increase their assortment size when faced with more intense competition: as the distance to the closest competitor increases by 1km, assortment sizes decrease by 7.99 products, or about 9% decrease from the average.[14] This relationship is similar for both independent retailers and chains (Appendix A.8).

We next examine how competition affects assortment composition. We measure assortment differentiation using the pairwise similarity measure of assortment overlap described in Section 2.3. For each pair of retailers, the variable captures the overlap in the identities of the supplied products. This differentiation variable is pairwise in nature, so we focus on understanding the relationship between distance to the nearest competitor and the assortment similarity with this closest competitor.[15] The results in Column (1) showed that retailers strategically adjust their assortment sizes in response to competition, so we control for the size of assortment in the subsequent analyses.

Column (2) reports a significant positive coefficient on the distance to the nearest competitor, indicating that retailers located farther apart supply more similar product assortments. This pattern runs counter to what one might expect if customer preferences and distribution costs were the dominant forces shaping assortments. For example, one might expect that customers near each other are likely to have similar tastes, suggesting that stores in closer geographic proximity face similar local demand. Similarly, producers may benefit from economies of scale when distributing to multiple stores in the same area. While these effects push in the direction of higher assortment similarity among closer competitors, we find that these pressures are offset by strategic differentiation motives.

To better understand the nature of assortment differentiation, we explore whether retailers differentiate horizontally or vertically. Horizontal differentiation refers to variation in product offerings that cater to different consumer tastes and preferences, while vertical differentiation involves differences in product quality, where higher quality usually commands a higher price. Distinguishing between these two types of differentiation provides deeper insight into the competitive strategies and market positioning.

We construct two measures of horizontal differentiation. The first tracks differentiation at the product category level and asks whether retailers in close proximity differentiate themselves by emphasizing different categories, for instance, by specializing in edible products vs. extract products. As most retailers supply all product categories, we operationalize this idea using the retailer's share of product items offered that are supplied in each category. We then construct the cosine similarity between the focal retailer and its closest competitor using these shares. This measure allows us to

[14] The direction of endogeneity bias is ambiguous for assortment size, as both store scale and local demand conditions affect the number of products carried. We report the naïve OLS results in Appendix A.8.

[15] Because the similarity measure is defined at the retailer pair level, we construct a single control for expected distance to the nearest competitor by averaging the expected distance to the nearest competitor across the two retailers in the pair.

quantify the degree of differentiation in product categories between retailers. Column (3) of Table 10 reports a positive coefficient, indicating that nearby competitors differentiate by specializing in different product categories.

The second measure exploits the fact that usable (flower) marijuana is the most homogeneous and least brand-differentiated category. We speculate that a higher share of usable products offers less scope for horizontal differentiation. Column (4) shows that proximity to a competitor reduces the share of usable product items, again consistent with greater horizontal differentiation among closely located firms.

Finally, we examine whether the observed reduction in assortment overlap between nearby competitors also reflects vertical differentiation. To proxy for the quality of a retailer's assortment, we calculate the (simple) average wholesale price of products in its assortment. We then compute the absolute difference in average wholesale prices between each retailer and its nearest competitor. Larger differences would suggest greater vertical differentiation, consistent with retailers offering products of systematically higher or lower quality. The results, shown in column (5) of Table 10, reveal no statistically significant relationship between this variable and proximity to competitors.

Taken together, the evidence in Table 10 suggests that the differentiation observed in the data is likely driven by horizontal differentiation, with little support for vertical differentiation. Retailers appear to tailor their assortments by emphasizing different product categories or moving away from the more commodified usable product category, rather than by systematically adjusting product quality levels.

6 Competition, Differentiation, and Markups

The analyses above suggest that retailers differentiate themselves in response to competition by supplying larger and more differentiated product assortments. Theoretical predictions regarding the consequences of such differentiation for pricing are mixed. On one hand, Anderson and Renault (1999) show that increased differentiation may induce more consumer search, thereby intensifying price competition and reducing markups. By contrast, Chen and Riordan (2008) show that if the differentiated products appeal to distinct market segments, it would steepen both firms' residual demand, leading to higher prices. Whether and to what extent differentiation is successful in softening price competition is an empirical question, which we analyze next.

To evaluate this question, we distinguish between store-level competition and product-level competition. Store-level competition captures the number of nearby retailers regardless of what they sell, while product-level competition captures the number of those retailers that carry the same product j. If differentiation softens competition, then the product-level effect should capture the additional pressure that arises when competitors sell the same product, whereas the store-level effect should capture a general competitive pressure on all of a retailer's products. We evaluate this prediction by looking at competition at the product level, which we define as the number of nearby competitors supplying each product within 0.5 and 1km radii. We incorporate this measure into our main markup specification from Equation 5, estimating:

$$\text{Markup}_{jrmt} = \alpha \text{ProdComp}_{jrt} + \beta \text{Comp}_r + \delta \mu_r + \eta_{jm} + \theta_{jt} + \varepsilon_{jrmt}. \tag{14}$$

where Markup_{jrmt} is again the markup for product j sold by retailer r in city m and quarter t. The variable ProdComp_{jrt} tracks the number of retailers supplying product j within 0.5km or 1km radii, while Comp_r captures the overall number of nearby competitors regardless of product. As in previous sections, we control for the expected number of competitors, μ_r, and include city-by-product fixed effects, η_{jm}, and quarter-by-product fixed effects, θ_{jt}, to account for time-varying product demand shocks and market-level product preferences.

A potential concern with equation (14) is that ProdComp_{jrt} may be endogenously correlated with unobserved local variation in consumer demand at a finer spatial level than is captured by the city-by-product fixed effects. Demand in local neighborhoods can differ in price sensitivity, search behavior, or preference for variety, and these differences may simultaneously affect the assortments and markups. For example, a neighborhood with more price-shopping consumers could be characterized by lower markups and higher retailer differentiation. To evaluate the extent to which this concern may bias our estimates, we estimate an additional specification that includes retailer fixed effects:

$$\text{Markup}_{jrmt} = \alpha \text{ProdComp}_{jrt} + \eta_{jm} + \theta_{jt} + \psi_r + \varepsilon_{jrmt}. \tag{15}$$

The inclusion of ψ_r absorbs all time-invariant characteristics of each retailer and its local market, including managerial quality, pricing strategies, brand reputation, and persistent differences in local demand. A direct consequence of this specification is that we can no longer identify the effect of overall store-level competition, β, as this variable is absorbed by the retailer fixed effects. The identifying assumption in equation (15) is that any relevant heterogeneity in consumer demand that might be correlated with assortments and markups is time-invariant.[16]

Columns (2) and (5) of Table 11 present our main results, which include both the number of stores in the area and the number of competitors at the product level. We find that product-level competition does have a direct impact on markups beyond our general measure of local competition for the store. That is, if the retailer is the only supplier of a product in the local area, the retailer can charge higher prices for that product. Importantly, the store-level competition coefficient remains economically meaningful even after controlling for product-level overlap, indicating that both margins matter. For example, the result in column (2) suggests that relative to a local monopolist, a pair of duopolists selling the same product would set markups \$0.65 (0.41+0.24) lower. If those duopolists sold differentiated products, by contrast, they would set markups only \$0.41 lower. This difference is consistent with assortment differentiation softening the intensity of price competition

[16] The lottery does not provide a direct instrument for product-level competition because the decision of whether to carry a product is itself an endogenous response to lottery outcomes, not a pre-determined characteristic of the applicant.

Table 11: Effect of *Product Level* Competition on Markups

	(1)	(2)	(3)	(4)	(5)	(6)
# Comp (0.5km)	–0.454***	–0.409***				
	(0.053)	(0.051)				
E[# Comp] (0.5km)	0.079*	0.075				
	(0.047)	(0.046)				
# Comp. by Prod (0.5km)		–0.237***	–0.203***			
		(0.052)	(0.033)			
# Comp (1km)				–0.298***	–0.272***	
				(0.035)	(0.033)	
E[# Comp] (1km)				0.062*	0.058	
				(0.038)	(0.038)	
# Comp. by Prod. (1km)					–0.141***	–0.157***
					(0.035)	(0.023)
N	194,509	194,509	194,509	194,509	194,509	194,509
City-product FE	yes	yes	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes	yes	yes
Retailer FE	no	no	yes	no	no	yes

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Columns (1)–(3) measure product-level competition within 0.5km; columns (4)–(6) within 1km. Columns (3) and (6) add FE at the retailer level. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

between nearby retailers.

To assess whether this product-level competition effect could be confounded by persistent local demand differences, such as variation in price sensitivity across neighborhoods, we include retailer fixed effects in columns (3) and (6). These absorb time-invariant characteristics of each store and its surrounding customer base. The coefficient on product-level competition remains stable, suggesting that the main results from equation (14) are not meaningfully driven by bias from unobserved local demand variation.

7 Evaluating the Impact of Entry Caps

Retail license caps are a common policy tool in newly legalized cannabis markets, but their effects on market outcomes remain underexplored.[17] The preceding analyses imply that the retail license cap created rents for lottery winners by restricting competition, which effectively transferred surplus from consumers to retailers through higher markups. To assess the magnitude of these rents, we turn to a simulation exercise that examines the license cap design.

While the lottery provides clean identification of competitive effects without the need for a

[17] A total of 15 U.S. states and several Canadian provinces currently employ similar license caps in the cannabis industry. Similar regulations are also used for alcohol or tobacco retail licenses in 17 states. While less common, some states have regulations restricting the number of pharmacies, gas stations, and lottery outlets. A closely related policy is dispersal regulations such as those restricting entry by dollar stores in many cities (Caoui, Hollenbeck, and Osborne, 2026).

structural model, this also limits our ability to perform a full equilibrium counterfactual analysis without imposing ad hoc assumptions on consumer demand and firm strategy. Accordingly, we limit our attention to outcomes that are cleanly identified in the preceding analyses, namely the causal effect of additional competitors on retailers' markups and product assortments.[18] As such, the simulations below constitute a *partial equilibrium comparative static*: we hold quantity demanded fixed and do not attempt to model general equilibrium responses, such as how the new entrants would set prices or decide their assortments. Nevertheless, this exercise provides informative and policy-relevant lower bounds on the likely impact of relaxing entry caps.

A key insight from our empirical findings motivating this exercise is that retail competition in this market is highly local; therefore, the effect of an additional entrant depends on where that entry occurs. This highlights again the importance of endogenous entry: we need to account for both the number of new entrants under a more permissive regime as well as where those entrants are expected to locate.

In our simulation, we increase the cap on the number of stores in each lottery jurisdiction by 10% and compute counterfactual outcomes for each observed retailer under the new market structure.[19] This corresponds to 84 additional stores across the state. The number of new entrants varies proportionally with market size. For instance, the two largest cities, Seattle and Spokane, account for the bulk of additional entry with 30 new stores.

We simulate 500 counterfactual lottery draws under the expanded license cap, holding fixed the original lottery winners and their locations. For each draw, we identify which additional applicants would have won and where their stores would have been located, assuming all new lottery winners enter the market. These simulations draw from the true set of applicants and their proposed locations. We then recompute the competitive environment around each existing retailer, focusing on the number of competitors within a 0.5 km radius, and apply our estimated coefficient of -0.454 (shown in column (2) of Table 8) to predict the resulting change in markups. This yields expected markup changes under the counterfactual policy. We also compute consumer expenditures under the assumption that the total quantity sold is fixed.

Table 12 summarizes the simulation results, with average changes calculated at the jurisdiction level. The simulation results show that an increase in the retail license cap by 10% leads to a decrease in markups of $0.11 on average, or by 0.7% overall. While this is the average for all stores, the price decrease varies substantially across stores based on the presence of nearby applicants (and hence the increase in competition from relaxing the license cap). The effects vary across jurisdictions based on the degree of excess demand for licenses by potential entrants, which we capture using the lottery winning probability. Price decreases are largest in jurisdictions with low winning probabilities, where excess demand for entry is greatest, and smallest in markets where most applicants already

[18] Thomas (2019) and Pavlov (2025) study aspects of the entry lottery using structural analyses, with Thomas (2019) focusing on the license cap and Pavlov (2025) focusing on the use of lottery allocation rather than an auction.

[19] In jurisdictions where the 10% increase in stores is less than 1 store, we increase the cap by a minimum of 1 store. Therefore, while most jurisdictions, and all large jurisdictions, experience a 10% increase, the average increase statewide is higher than 10%.

Table 12: Market Level Outcomes

	Δ Avg Price ($)	Δ Avg Assort. Size	Annual Consumer Savings
All Markets	-.109	19.5	$7,507,000
$P(win) < \frac{1}{3}$	-.143	25.7	$6,010,000
$\frac{1}{3} < P(win) < \frac{2}{3}$	-.075	13.6	$1,404,000
$P(win) > \frac{2}{3}$	-.055	9.1	$92,000
Seattle	-.250	44.9	$2,253,000
Rest of State	-.089	16.0	$5,253,000

Notes: The table reports average changes from a simulated 10% increase in the retail license cap in each lottery jurisdiction. Results are averaged across 500 counterfactual lottery draws. Markup changes are in dollars. Consumer savings are annual, computed holding total quantity sold fixed.

won licenses. Comparing the largest urban market, Seattle, to the rest of the state, we find that loosening the license cap by 10% would reduce markups by $0.25 on average in Seattle, due to the much higher density of the retail market.

In addition to lowering markups, increasing the cap would increase product variety available to consumers. Existing retailers increase their assortment size by roughly 20 product items on average, or 8.4%. We note that this should be seen as a lower bound on the increase in product variety, as it only considers the impact on the existing set of retailers, and the hypothetical entrants in this scenario would further increase product variety.

The last column in Table 12 reports estimated annual consumer savings from lower prices, holding total quantity sold fixed. Overall, relaxing the license cap by 10% would generate roughly $7.5 million per year in consumer savings. This suggests that consumers are modestly but meaningfully harmed by the state's retail license cap, although we note that this estimate likely understates the true consumer benefit, as it captures only the price channel and omits gains from increased product variety, reduced travel costs, and the additional assortment that new entrants would bring.

Lastly, we demonstrate the importance of considering spatial market structure and endogenous entry locations in these simulations, relative to simply extrapolating from the results estimated in section 4. To do so, we make two additional comparisons. First, we compare our results to a naïve benchmark that does not account for endogenous entry locations. Instead of drawing new entrants from the actual applicant pool, we compute each store's expected change in competition under the assumption that new entrants locate uniformly at random within the market. This yields an average price decrease of only $0.05, compared with $0.11 in our main simulation. The smaller decrease in prices reflects the fact that relaxing the entry cap would disproportionately generate entry near existing stores. Failing to account for endogenous entry locations, in other words, would lead us to underestimate the impact of the policy change by half.

Second, we compare the distribution of market-level price changes across our 500 simulated lottery draws. While we focus on the average price change across simulations above, there is substantial heterogeneity across lottery draws in how much prices would change, and we highlight this heterogeneity to emphasize again the importance of the spatial distribution of entry. To do so, we compute the market-level price change for each of the 500 simulations and then rank them from largest to smallest for each market. We then take the average of the smallest and largest price changes across all cities. While the average expected decrease in prices is \$0.11, we find that, depending on the spatial configuration of new entrants, simulated outcomes range from only a \$0.03 decrease to a \$0.23 decrease in average markups, highlighting the importance of entry location in the policy implications studied.

The primary policy implication is that entry restrictions substantially harm consumers and benefit the arbitrary subset of retail applicants who won store licenses in the lottery, but there are several other implications as well. First, the retail license cap also benefits retailers at the expense of firms operating in the upstream market. Product manufacturers (and farmers producing the initial plants) both face lower total demand due to higher retail markups and experience less demand for product variety and smaller downstream assortments. This second aspect likely means that the upstream market is both less profitable and more concentrated as a result. In addition, if it is a policy goal of the state to restrict demand and keep prices high to limit cannabis consumption, doing so via restrictions on retail competition is not an efficient mechanism. This can be seen via the high rents going directly to retailers, where an alternative policy instrument, like excise taxes, could accomplish the same goal but with the benefit going to taxpayers instead. That is, the state could allow more retail competition and then offset the decrease in prices with higher taxes, therefore capturing the rents currently earned by lottery winners.

8 Conclusion

The effect of competition on retail markups and product assortment decisions is a central question in industrial organization and marketing, with direct implications for merger analysis and entry regulation. In this paper, we combine granular transaction-level data on retail sales and prices with random variation in local market structure to estimate the causal effects of competition on retail markups and assortments.

The unique feature of our setting, the legal cannabis retail sector in Washington State, is the lottery allocation of store licenses following an open application process. Applicants had to pre-commit to store locations prior to the lottery; therefore, the number and spatial distribution of local competitors were randomly assigned among applicants. However, we still face an identification challenge as the set of proposed locations is itself endogenous. Mainly, firms likely selected locations they expected to be profitable, thus, areas with more applicants may differ systematically. Our identification strategy, following Borusyak and Hull (2023), exploits the fact that conditional on the pre-lottery application environment, the realized level of competition varies randomly due to the lottery. We construct a measure of expected competition for each retailer based on the number and proposed locations of nearby applicants and the city-level winning probability, and control for this measure in our regressions. The coefficient on realized competition then reflects only the lottery-driven deviations from expected competition, which are orthogonal to the endogenous location choices. Overall, the empirical setting gives us rich variation in competitive structure and unusually detailed data on retailer markups and assortments, allowing us to trace out how competition shapes markups and assortment decisions.

There are a number of key findings. First, we show that nearby competition significantly reduces markups, but with strongly diminishing returns as the number of local competitors increases. Moving from a monopoly to two or three nearby competitors leads to a large decrease in markups, after which the presence of additional competitors has little further impact. We also find that independent retailers respond more strongly to local competition than chains. Second, retailers facing more nearby competition carry more differentiated assortments. The closer a competitor is, the more likely a given retailer will choose an assortment with less overlap. Third, we document that competition affects markups both broadly, at the store level, and narrowly, at the specific product level. Product-level overlap between neighboring retailers leads to additional markup reductions beyond store-level competition, consistent with assortment differentiation weakening price competition. These findings present an interesting tradeoff from the perspective of consumers. As competitive pressure in a market increases, firms may react by increasing their product differentiation in order to maintain higher prices, which have opposing effects on consumer welfare. Consumers benefit from the increase in variety this produces, but may be worse off if this causes prices to be substantially higher as a result. We leave the formal evaluation of this tradeoff as the subject of future work.

Finally, we simulate the effect of relaxing the state's retail license cap. Using our causal estimates, we model a 10% increase in licenses, drawing from the observed applicant pool and their proposed locations. This partial equilibrium exercise (holding quantities sold constant) shows that such a policy would generate roughly $7.5 million in annual consumer savings through lower prices, while also increasing product variety. Effects are strongest in markets with high excess demand for entry. These gains are likely conservative, as we do not account for new entrant variety or expansion of market demand. The simulation highlights how retail license caps transfer rents to incumbents and reduce consumer welfare.

References

AILAWADI, K. L., AND K. L. KELLER (2004): "Understanding retail branding: conceptual insights and research priorities," *Journal of Retailing*, 80(4), 331–342.

AILAWADI, K. L., J. ZHANG, A. KRISHNA, AND M. W. KRUGER (2010): "When Wal-Mart enters: How incumbent retailers react and how this affects their sales outcomes," *Journal of Marketing Research*, 47(4), 577–593.

ANDERSON, S. P., AND R. RENAULT (1999): "Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model," *The RAND Journal of Economics*, 30(4), 719–735.

ARCIDIACONO, P., P. B. ELLICKSON, C. F. MELA, AND J. D. SINGLETON (2020): "The Competitive Effects of Entry: Evidence from Supercenter Expansion," *American Economic Journal: Applied Economics*, 12(3), 175–206.

ARMSTRONG, M., AND J. VICKERS (2022): "Patterns of Competitive Interaction," *Econometrica*, 90, 153–191.

ATKIN, D., B. FABER, AND M. GONZALEZ-NAVARRO (2018): "Retail Globalization and Household Welfare: Evidence from Mexico," *Journal of Political Economy*, 126(1), 1–73.

BAUNER, C., AND E. WANG (2019): "The Effect of Competition on Pricing and Product Positioning: Evidence from Wholesale Club Entry," *International Journal of Industrial Organization*, 67, 102525.

BERRY, S., A. EIZENBERG, AND J. WALDFOGEL (2016): "Optimal Product Variety in Radio Markets," *The RAND Journal of Economics*, 47(3), 463–497.

BOLTON, R., AND V. SHANKAR (2003): "An empirically derived taxonomy of retailer pricing and promotion strategies," *Journal of Retailing*, 79, 213–224.

——— (2004): "An Empirical Analysis of Determinants of Retailer Pricing Strategy," *Marketing Science*, 23, 28–49.

BONFRER, A., P. CHINTAGUNTA, AND S. DHAR (2022): "Retail store formats, competition and shopper behavior: A Systematic review," *Journal of Retailing*, 98(1), 71–91.

BORUSYAK, K., AND P. HULL (2023): "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 91(6), 2155–2185.

BRESNAHAN, T. F., AND P. C. REISS (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), 977–1009.

BRIESCH, R. A., P. K. CHINTAGUNTA, AND E. J. FOX (2009): "How Does Assortment Affect Grocery Store Choice?," *Journal of Marketing Research*, 46(2), 176–189.

BRIESCH, R. A., W. R. DILLON, AND E. J. FOX (2013): "Category Positioning and Store Choice: The Role of Destination Categories," *Marketing Science*, 32(3), 488–509.

BRONIARCZYK, S. M., W. D. HOYER, AND L. MCALISTER (1998): "Consumers' perceptions of the assortment offered in a grocery category: The impact of item reduction," *Journal of Marketing Research*, 35(2), 166–176.

Bronnenberg, B. J., J.-P. Dubé, and S. Moorthy (2019): "The Economics of Brands and Branding," *Handbooks in Economics; Economics of Marketing*, 1.

Busso, M., and S. Galiani (2019): "The Causal Effect of Competition on Prices and Quality: Evidence from a Field Experiment," *American Economic Journal: Applied Economics*, 11(1), 33–56.

Caoui, E. H., B. Hollenbeck, and M. Osborne (2026): "Dynamic Entry Spatial Competition: An Application to Dollar Store Expansion," *RAND Journal of Economics*.

Caulkins, J., and L. Dahlkember (2013): "Retail Store Allocation," I-502 project 430, BOTEC Analysis Corp.

Chan, T. Y., V. Padmanabhan, and P. Seetharaman (2007a): "An Econometric Model of Location and Pricing in the Gasoline Market," *Journal of Marketing Research*, 44(4), 622–635.

——— (2007b): "An Econometric Model of Location and Pricing in the Gasoline Market," *Journal of Marketing Research*, 44(4), 622–635.

Chen, Y., and M. H. Riordan (2008): "Price-Increasing Competition," *The RAND Journal of Economics*, 39(4), 1042–1058.

Cirik, K., and R. Makadok (2023): "Online Reviews, Market Rivalry, and Pioneer Advantage: Evidence from a Natural Experiment with Randomized Entry Order in Marijuana Retailing," *Academy of Management Review*, 48(3), 409–431.

City of Tacoma (2014): "Initiative 502 Frequently Asked Questions," `https://cms.cityoftacoma.org/planning/recreational%20marijuana/I-502%20FAQs.pdf`, Accessed February 23, 2026.

Conlon, C. T., and J. H. Mortimer (2021): "Efficiency and Foreclosure Effects of Vertical Rebates: Empirical Evidence," *Journal of Political Economy*, 129(12), 3357–3404.

d'Aspremont, C., J. Gabszewicz, and J. Thisse (1979): "On Hotelling's "Stability in Competition"," *Econometrica*, 47(5), 1145–50.

Datta, S., and K. Sudhir (2013): "Does reducing spatial differentiation increase product differentiation? Effects of zoning on retail entry and format variety," *Quantitative Marketing and Economics (QME)*, 11(1), 83–116.

Davis, L., S. McRae, and E. Seira (2025): "Competitive Effects of Entry in Gasoline Markets," *Journal of Industrial Economics*, 73(4).

DellaVigna, S., and M. Gentzkow (2019): "Uniform pricing in us retail chains," *The Quarterly Journal of Economics*, 134(4), 2011–2084.

Draganska, M., M. Mazzeo, and K. Seim (2009): "Beyond plain vanilla: modeling joint product assortment and pricing decisions," *Quantitative Marketing and Economics*, 7(2), 105–146.

Dukes, A. J., T. Geylani, and K. Srinivasan (2009): "Strategic Assortment Reduction by a Dominant Retailer," *Marketing Science*, 28(2), 309–319.

Eizenberg, A. (2014): "Upstream Innovation and Product Variety in the U.S. Home PC Market," *The Review of Economic Studies*, 81, 1003–1045.

Fan, Y., and C. Yang (2020): "Competition, Product Proliferation, and Welfare: A Study of the US Smartphone Market," *American Economic Journal: Microeconomics*, 12(2), 99–134.

Fischer, K., S. Martin, and P. Schmidt-Dengler (2025): "The Heterogeneous Effects of Entry on Prices," *Journal of the European Economic Association*.

Fox, E. J., A. L. Montgomery, and L. M. Lodish (2004): "Consumer Shopping and Spending across Retail Formats," *The Journal of Business*, 77(S2), S25–S60.

Ginsburgh, V., A. de Palma, Y. Papageorgiou, and J. Thisse (1985): "The Principle of Minimum Differentiation Holds under Sufficient Heterogeneity," *Econometrica*, 53(4), 767–81.

Hansen, B., K. Miller, and C. Weber (2020): "Federalism, Partial Prohibition, and Cross Border Sales: Evidence from Recreational Marijuana," *Journal of Public Economics*, 187, 104–159.

Hitsch, G. J., A. Hortaçsu, and X. Lin (2021): "Prices and promotions in U.S. retail markets," *Quantitative Marketing and Economics*, 19(3-4), 299–451.

Hoch, S. J., E. T. Bradlow, and B. Wansink (1999): "The Variety of an Assortment," *Marketing Science*, 18(4), 527–546.

Hollenbeck, B. (2017): "The Economic Advantages of Chain Affiliation," *RAND Journal of Economics*, 48, 1103–1135.

Hollenbeck, B., and R. Z. Giroldo (2021): "Winning Big: Scale and Success in Retail Entrepreneurship," *Marketing Science*, 41(2), 271–293.

Hollenbeck, B., and K. Uetake (2021): "Taxation and Market Power in the Legal Marijuana Industry," *RAND Journal of Economics*, 52, 559–595.

Hosken, D. S., R. S. McMillan, and C. T. Taylor (2008): "Retail gasoline pricing: What do we know?," *International Journal of Industrial Organization*, 26(6), 1425–1436.

Hristakeva, S. (2022a): "Determinants of Channel Profitability: Retailers' Control over Product Selections as Contracting Leverage," *Marketing Science*, 41(2), 315–335.

——— (2022b): "Vertical Contracts with Endogenous Product Selection: An Empirical Analysis of Vendor Allowance Contracts," *Journal of Political Economy*, 130(12), 3202 – 3252.

Hwang, M., B. J. Bronnenberg, and R. Thomadsen (2010): "An Empirical Analysis of Assortment Similarities across US Supermarkets," *Marketing Science*, 29(5), 858–879.

Ilanes, G., and S. Moshary (2019): "Market Structure and Product Assortment: Evidence from a Natural Experiment in Liquor Licensure," *Kilts Center for Marketing at Chicago Booth – Nielsen Dataset Paper Series 2-013.*

Krishan, T., K. Koelemeijer, and R. Rao (2002): "Consistent assortment provision and service provision in a retail environment," *Marketing Science*, 21(1), 54–73.

Luo, B. (2024): "Manufacturer–Retailer Relationships and the Distribution of New Products," *Marketing Science*, 43(1), 114–137.

Matsa, D. A. (2011): "Competition and Product Quality in the Supermarket Industry*," *The Quarterly Journal of Economics*, 126(3), 1539–1591.

Mazzeo, M. J. (2002): "Competitive Outcomes in Product-Differentiated Oligopoly," *The Review of Economics and Statistics*, 84(4), 716–728.

Mogilner, C., T. Rudnick, and S. S. Iyengar (2008): "The Mere Categorization Effect: How the Presence of Categories Increases Choosers' Perceptions of Assortment Variety and Outcome Satisfaction," *Journal of Consumer Research*, 35(2), 202–215.

Oschmann, S. (2023): "The Competitive Effects of Superstore Entry: The Role of Product Variety," *mimeo*.

Pancras, J., S. Sriram, and V. Kumar (2012): "Empirical investigation of retail expansion and cannibalization in a dynamic environment," *Management Science*, 58(11), 2001–2018.

Pavlov, V. (2025): "The Price of Fairness: Equity and Efficiency in Retail License Allocation," *Working Paper*.

Pozzi, A., and F. Schivardi (2016): "Entry Regulation in Retail Markets," *Handbook on the Economics of Retailing and Distribution*, 1.

Ren, C. R., Y. Hu, Y. J. Hu, and J. Hausman (2011): "Managing Product Variety and Collocation in a Competitive Environment: An Empirical Investigation of Consumer Electronics Retailing," *Management Science*, 57(6), 1009–1024.

Rosenthal, R. W. (1980): "A Model in which an Increase in the Number of Sellers leads to a Higher Price," *Econometrica*, 48(6), 1575–1579.

Satterthwaite, M. A. (1980): "Consumer Information, Equilibrium Industry Price, and the Number of Sellers," *The Bell Journal of Economics*, 10(2), 483–502.

Singh, V., and T. Zhu (2008): "Pricing and Market Concentration in Oligopoly Markets," *Marketing Science*, 27(6), 1020–1035.

Stiglitz, J. E. (1987): "Competition and the Number of Firms in a Market: Are Duopolies More Competitive than Atomistic Markets?," *Journal of Political Economy*, 95(5), 1041–1061.

Sweeting, A. (2010): "The Effects of Mergers on Product Positioning: Evidence from the Music Radio Industry," *The RAND Journal of Economics*, 41(2), 372–397.

Thomadsen, R. (2005): "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," *RAND Journal of Economics*, 36, 908–929.

Thomadsen, R. (2007): "Product Positioning and Competition: The Role of Location in the Fast Food Industry," *Marketing Science*, 26(6), 792–804.

Thomas, D. (2019): "License Quotas and the Inefficient Regulation of Sin Goods: Evidence from the Washington Recreational Marijuana Market," *Working Paper*.

Tirole, J. (1988): *The Theory of Industrial Organization*, vol. 1 of *MIT Press Books*. The MIT Press.

Watson, R. (2009): "Product Variety and Competition in the Retail Market for Eyeglasses," *Journal of Industrial Economics*, 57(2), 217–251.

Wollmann, T. G. (2018): "Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 108(6), 1364–1406.

Zhu, T., V. Singh, and A. Dukes (2011): "Local competition, entry, and agglomeration," *Quantitative Marketing and Economics*, 9(2), 129–154.

A Appendix: Additional Results

A.1 Selection Tests

In this section, we test for potential selection on observable retailer characteristics. A potential concern is that different entry locations attract different retailer types. While the lottery provides clean identification of the causal effect of number of competitors on outcomes like markups, the interpretation of these effects would be different if locations with more/fewer competitors were systematically of different types. An example would be if high demand locations attracted high-quality applicants, i.e. those with greater capital or managerial ability. We therefore regress observable characteristics of applicants on the local competition variables and show results in Table A.1. The characteristics we study are the total number of applications by each retail firm and the number of licenses won (chain size). Total number of applications is constructed at the firm level using compiled data across all store applications, which are linked based on a unique firm identifier. These include applications across all markets, including both lottery and non-lottery markets.

In addition, prior to the passage of I-502 which created the legal cannabis industry, there existed a decriminalized system of medical dispensaries where consumers with a doctor's prescription could legally buy cannabis products, although the sellers of those products existed in a legal grey area. Those dispensaries may have secured prime retail locations and may have built up consumer relationships or industry experience that would be valuable in the post-legalization market. To capture which applicants were medical dispensaries, we hand-collect data from the Internet Archive stored records of weedmaps.com as they existed in 2014. This website provided store location information which we use to match post-legalization stores to pre-legalization stores. We use this to construct a dummy for whether the store or firm existed in the pre-legalization market as a medical dispensary, which is a proxy for industry experience.

We find few significant relationships. Of 24 coefficients estimated, only 2 are significant at the 5% level. Both are for the number of applications, in which we find a significant relationship between the number of applications and the likelihood of being in a location with only 1 competitor. This could be a spurious result or could mechanically suggest that firms with more applications are more likely to win licenses in areas with fewer competitors. Importantly, none of the coefficients on chain size or former medical dispensary status are significant, suggesting that the type of retailer operating in a location is not systematically related to the realized level of competition.

Table A.1: Test of Selection on Retailer Characteristics

	(1) # Stores in Chain	(2) # Applications	(3) Former Medical	(4) # Stores in Chain	(5) # Applications	(6) Former Medical
# Comp. (1km)=1	0.26	0.49**	–0.05			
	(0.17)	(0.19)	(0.04)			
# Comp. (1km)=2	–0.05	0.05	0.00			
	(0.20)	(0.23)	(0.05)			
# Comp. (1km)=3	0.22	0.07	0.09			
	(0.30)	(0.34)	(0.07)			
# Comp. (1km)=4	0.09	0.24	0.05			
	(0.27)	(0.30)	(0.06)			
# Comp. (0.5km)=1				0.19	0.37**	–0.02
				(0.16)	(0.18)	(0.04)
# Comp. (0.5km)=2				–0.18	0.08	–0.03
				(0.22)	(0.25)	(0.05)
# Comp. (0.5km)=3				0.23	0.15	0.09
				(0.39)	(0.46)	(0.10)
# Comp. (0.5km)=4				0.22	0.83	0.02
				(0.49)	(0.52)	(0.11)
N	397	341	341	397	341	341

Notes: The unit of observation is at the retailer level. The dependent variables are retailer characteristics measured prior to market entry. Columns (1)-(3) measure competition within a 1km radius; columns (4)-(6) within a 0.5km radius. The sample size differs across columns because the number of applications and former medical dispensary status could not be determined for all retailers, while chain size is observable for the full sample. Standard errors are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.2 Alternative Functional Form

Below, we use an alternative strategy that isolates variation in realized competition within more homogenous competitive environments. Specifically, we estimate the effect of realized competition on markups separately for subsamples defined by the expected number of nearby competitors. That is, for each rounded value of expected competition within 1km, we run separate regressions of markups on the realized number of competitors. These specifications allow us to compare firms that faced the same pre-lottery application environment but differ only in the number of competitors they ended up facing due to the lottery outcome.

The estimates in Table A.2 show consistent negative effects of realized competition on markups across bins. While the magnitude of these effects varies somewhat across expected competition levels, the pattern is not systematic. This strengthens our interpretation that the observed relationship between competition and markups is not driven by differences in entrant types across markets, but reflects genuine competitive effects.

Using this non-parametric approach more centrally in the paper is tempting as it relies on weaker functional form assumptions and is robust to selection across application environments. However, estimating treatment effects separately within each expected competition bin reduces the sample size in each regression and limits statistical power. For this reason, we adopt the pooled specification as our preferred approach in the main analyses, while using the non-parametric specification as a robustness check that confirms the validity of our identification strategy.

Table A.2: Effects of Retailer Competition on Markups: By Expected Market Competitive Pressures

	(1) E[# Comp]=0	(2) E[# Comp]=1	(3) E[# Comp]=2	(4) E[# Comp]$\geq$ 3
# Comp (1km)	–0.369***	–0.214**	–0.309**	–0.152***
	(0.076)	(0.090)	(0.151)	(0.031)
N	90,614	66,287	18,027	19,581
City-product FE	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Each column restricts the sample to retailers whose expected number of competitors within 1km (rounded) equals the value indicated in the column header. All regressions include FE for quarter-by-product and city-by-product. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.3 First Stage Results

As discussed in Section 3.1, our preferred estimator is the controlled OLS specification following Borusyak and Hull (2023), but we also report IV estimates in the main tables for comparison. Table A.3 shows the first stage regressions for these IV specifications, regressing realized competition on recentered competition. Expected competition is a strong predictor of realized competition across all three measures, with Kleibergen-Paap F statistics of 123.5, 46.04, and 121.33, respectively, well above conventional thresholds for instrument relevance. Partial R-squareds are .621, .152, and .201, respectively.

Table A.3: First Stage Results

	(1) Distance to Closest Comp	(2) # Comp (0.5km)	(3) # Comp (1km)
Recentered Dist. to Closest Comp.	0.618***		
	(0.056)		
Recentered # Comp (0.5km)		0.331***	
		(0.049)	
Recentered # Comp (1km)			0.510***
			(0.046)
N	194509	194509	194509
R^2	0.787	0.551	0.579
Adj. R^2	0.750	0.473	0.507
City-product FE	yes	yes	yes
Quarter-product FE	yes	yes	yes

Notes: The unit of observation is at the retailer-quarter level. The dependent variable in each column is the realized competition measure indicated in the column header; the independent variable is the corresponding recentered competition measure. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.4 Results using Travel Duration

In this section, we repeat the analyses from section 4, using travel duration to construct our measures of local competitive pressures. Travel time is computed using Google's API as the time it takes to drive from store A to store B on 07/17/2019 08:00 AM. Our measures of competition track the *travel* duration to the closest competitor, and the number of competitors within a 1 and 8-minute drive. Table A.4 shows the results when the dependent variable is markups (in $s). The main results remain unchanged.

Table A.4: Effect of Retail Competition on Markups

	(1)	(2)	(3)	(4)	(5)	(6)
Minutes to Closest Comp.	0.084***	0.160***				
	(0.019)	(0.041)				
# Comp. (1 mins)			–0.228***	–0.353***		
			(0.083)	(0.105)		
# Comp. (8 mins)					–0.066***	–0.102***
					(0.014)	(0.014)
E[Minutes to Closest Comp.]		–0.102**				
		(0.045)				
E[# Comp.] (1 mins)				0.135***		
				(0.045)		
E[# Comp.] (8 mins)						0.058***
						(0.016)
R^2	0.544	0.544	0.543	0.543	0.544	0.544
N	179,618	179,618	179,618	179,618	179,618	179,618
City-product FE	yes	yes	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes	yes	yes

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars. Competition is measured using travel duration: distance to closest competitor (minutes), number of competitors within a 1-minute drive, and number of competitors within an 8-minute drive. All regressions control for the expected value of the respective competition measure. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.5 Nonlinear Estimates

Table A.5 reports the full regression estimates underlying the main results in Figures 5 and 6 in the main text. Column (1) discretizes distance to the closest competitor into five bins: 0-0.2km, 0.2-0.4km, 0.4-1.1km, 1.1-2.5km, and >2.5km (omitted). Columns (2) and (3) use the number of competitors within 0.5km and 1km, respectively, discretized as 0 (omitted), 1, 2, 3, and 4+. Each specification includes dummies for the corresponding discretized expected competition measure as controls, following Equation 12. The coefficients on the expected competition dummies do not have a causal interpretation; they serve as controls that absorb endogenous variation in the applicant environment.

Table A.5: Nonlinear Effects of Competition on Markups: Full Regression Outputs

	Dist. to the Closest Comp.	# Comp. 0.5km	# Comp. 1km
Realized competition			
Bin 1	–0.713***	–1.037***	–0.696***
	(0.238)	(0.193)	(0.231)
Bin 2	–0.700***	–1.600***	–1.416***
	(0.234)	(0.220)	(0.193)
Bin 3	–0.505	–1.392***	–1.377***
	(0.315)	(0.183)	(0.197)
Bin 4	0.274	–2.185***	–1.391***
	(0.172)	(0.267)	(0.201)
Expected competition			
Bin 1	–0.596***	0.436**	0.279
	(0.119)	(0.179)	(0.193)
Bin 2	–0.081	0.543***	0.067
	(0.165)	(0.145)	(0.147)
Bin 3	–0.215	0.344**	0.318**
	(0.272)	(0.171)	(0.139)
Bin 4	–0.600**	0.040	0.098
	(0.247)	(0.194)	(0.235)
City-product FE	yes	yes	yes
Quarter-product FE	yes	yes	yes

Notes: The dependent variable is retail markups measured in dollars. Each column corresponds to a different competition measure: distance to closest competitor (column 1), number of competitors within 0.5km (column 2), and number of competitors within 1km (column 3). Dummies for the discretized expected competition measure are included as controls; these coefficients do not have a causal interpretation. Standard errors clustered at the city-by-quarter level are reported in parentheses. These are the underlying estimates for Figures 5 and 6.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.6 Results excluding Seattle

In this section, we repeat the analyses from sections 4, 5, and 6 excluding observations from the Seattle market. Seattle is by far the largest and densest market in Washington. One might be concerned that our results are driven primarily by competitive dynamics in this single large urban market. The results below show that the main findings are robust to excluding Seattle: the estimated effects of competition on markups, assortment differentiation, and product-level competition are similar in sign, magnitude, and significance to those reported in the main tables.

Table A.6: Effect of Retail Competition on Markups: Excluding Seattle

	(1)	(2)	(3)
Dist. to Closest Comp.	0.365***		
	(0.076)		
# Comp (0.5km)		–0.493***	
		(0.087)	
# Comp (1km)			–0.341***
			(0.053)
E[Dist. to Closest Comp.]	–0.322***		
	(0.070)		
E[# Comp] (0.5km)		0.289***	
		(0.072)	
E[# Comp] (1km)			0.167***
			(0.058)
N	162,371	162,371	162,371
City-product FE	yes	yes	yes
Quarter-product FE	yes	yes	yes

Notes: The unit of observation is at the retailer-product-quarter level. The dependent variable is retailer markups measured in dollars. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Effect of Competition on Retail Differentiation: Excluding Seattle

	Assort Size (1)	Product Assort Diff (2)	Category Assort Diff (3)	Usable Share of Assort (4)	abs(Wholesale Price Diff) (5)
Dist. to Closest Comp.	–7.918***	0.007***	0.003**	0.006***	0.018
	(0.887)	(0.003)	(0.001)	(0.002)	(0.019)
E[Dist. to Closest Comp.]	7.254***	0.020***	0.002	–0.000	–0.095**
	(1.143)	(0.004)	(0.003)	(0.003)	(0.040)
Assort. Size (log)		0.050***	0.038***	0.015	–0.223***
		(0.005)	(0.006)	(0.011)	(0.074)
N	1,620	1,620	1,620	1,620	1,614
City FE	yes	yes	yes	yes	yes
Quarter FE	yes	yes	yes	yes	yes

Notes: The unit of observation is at the retailer-by-quarter level. In column (1), the dependent variable is the number of unique products carried. In columns (2)-(5), dependent variables are defined relative to the retailer's nearest competitor (see text). Columns (2)-(5) control for assortment size. Results include FE for city and quarter. Standard errors are clustered at the city-by-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Effect of *Product Level* Competition on Markups: Excluding Seattle

	(1)	(2)	(3)	(4)	(5)	(6)
# Comp (0.5km)	–0.493***	–0.435***				
	(0.087)	(0.085)				
E[# Comp] (0.5km)	0.289***	0.279***				
	(0.072)	(0.072)				
# Comp. by Prod (0.5km)		–0.261***	–0.168***			
		(0.068)	(0.040)			
# Comp (1km)				–0.341***	–0.310***	
				(0.053)	(0.052)	
E[# Comp] (1km)				0.167***	0.164***	
				(0.058)	(0.058)	
# Comp. by Prod. (1km)					–0.138***	–0.125***
					(0.050)	(0.031)
N	162,371	162,371	162,371	162,371	162,371	162,371
City-product FE	yes	yes	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes	yes	yes
Retailer FE	no	no	yes	no	no	yes

Notes: The unit of observation is at the product-by-retailer-by-quarter level. The dependent variable is retailer markups measured in dollars.Columns (3) and (6) add FE at the retailer level. Standard errors are clustered at the city-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.7 Indirect Competition

The above analyses viewed the area around each retailer as a "self-contained" market. However, one may expect that retailers outside the analyzed radius may influence the markups of the focal retailer through their competitive pressures on the rest of the local market. Suppose retailer A has only one competitor, B, in a 1km radius; while retailer B has both A and C as competitors in its local market. Thus, C would be an "indirect" competitor to A because it affects B's strategic behavior, and, hence, A's.

We evaluate the importance of these interactions by looking at the effects of marginal competitors that are further away from the focal retailer. That is, we start by replicating our result connecting markups to the number of competitors within a 1km distance in column (1) of Table A.9. Then in column (2), we add the number of additional competitors observed between the area of 1 to 2km radii. Column (3) expands the radius and adds the number of additional competitors within a 1 to 3-km radius. The estimates for the marginal competitive pressures are small and not statistically significant, leading us to conclude that indirect competition plays a limited role in our setting. These results are consistent with our distance results, which show that competitive effects attenuate sharply beyond 1km.

Table A.9: The Role of Indirect Competition

	(1)	(2)	(3)
# Comp (1km)	–0.298***	–0.330***	–0.332***
	(0.035)	(0.034)	(0.035)
# Comp. (1-2km)		–0.086	
		(0.065)	
# Comp. (1-3km)			–0.014
			(0.019)
E[# Comp] (1km)	0.062*	0.129***	0.114***
	(0.038)	(0.032)	(0.035)
E[# Comp.] (1-2km)		0.310***	
		(0.079)	
E[# Comp.] (1-3km)			0.211***
			(0.042)
N	194,509	194,509	194,509
City-product FE	yes	yes	yes
Quarter-product FE	yes	yes	yes

Notes: Observation is at the retailer-product-quarter level. The dependent variable is retailer markups measured in dollars. All regressions include city-by-product and quarter-by-product fixed effects. Standard errors are clustered at the city-by-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.8 Additional Results on Assortment Differentiation

For completeness, Table A.10 below replicates the analysis from Table 10 in the main text using a simple OLS specification.

Table A.10: Effect of Competition on Retail Differentiation: OLS

	Assort Size (1)	Product Assort Diff (2)	Category Assort Diff (3)	Usable Share of Assort (4)	abs(Wholesale Price Diff) (5)
Dist. to Closest Comp.	–2.710***	0.014***	0.002**	0.005***	–0.011
	(0.565)	(0.002)	(0.001)	(0.001)	(0.014)
N	1,935	1,935	1,935	1,935	1,929
City FE	yes	yes	yes	yes	yes
Quarter FE	yes	yes	yes	yes	yes

Notes: The unit of observation is at the retailer-by-quarter level. Results include FE for city and quarter. Standard errors are clustered at the city-by-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We also repeat the analysis of how assortment size responds to the distance to the closest competitor, estimating the regression separately for independent stores and chains. Table A.11 reports the results. We find no meaningful difference in responses across store types. The rest of the analyses in Table 10 rely on pairwise measures of assortment similarity, which are defined at the retailer-pair level. As a result, these outcomes cannot be cleanly split by store type. For that reason, we limit this heterogeneity analysis to assortment size.

Table A.11: Effect of Competition on Retail Differentiation: By Store Type

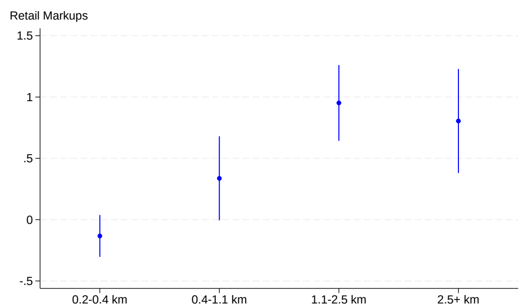
	Independent (1)	Chain (2)
Dist. to Closest Comp.	–6.969***	–6.746***
	(1.468)	(1.440)
E[Dist. to Closest Comp.]	6.490***	6.276***
	(1.492)	(1.603)
N	821	1,114
City FE	yes	yes
Quarter FE	yes	yes

Notes: The unit of observation is at the retailer-by-quarter level. Results include FE for the city and quarter. The regressions also include a dummy equal to 1 if the expected number of stores for the license retailer is greater than 1. Standard errors are clustered at the city-by-quarter level and reported under each parameter estimate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

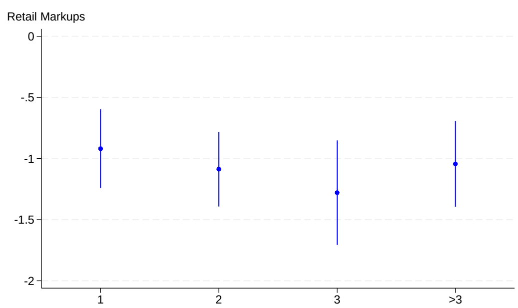
A.9 Matching Estimator

As a robustness test, we estimate the regressions from section 4 using a matching estimator. The results below use augmented IPW matching on the expected number of competitors. We impose exact matching on the quarter time periods. The main idea behind the estimator is to find stores with a similar expected number of competitors, which would capture differences in expected local demand and costs. Figure A.1 shows that the results under this alternative estimation strategy closely match our main findings in Figures 5 and 6.

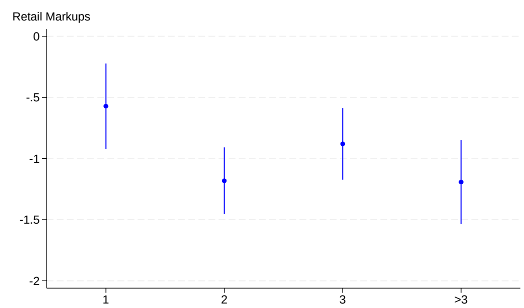
Figure A.1: Effects of Competition on Markup: Matching



(a) Distance to Closest Competitor



(b) Number of Competitors within 0.5km



(c) Number of Competitors within 1km

Notes: The unit of observation is at the product-retailer-quarter level. The dependent variable is retailer markups measured in dollars. Results are obtained using augmented ipw matching on the respective variable tracking the expected competitive pressures. The excluded category in panel (a) is retailers with a competitor within a 0-0.2km radius. In panels (b) and (c), the excluded category is defined as retailers without any competitors within 1.5 kilometers.

A.10 Distribution of Identifying Variation

As we noted in section 3, our estimator implicitly weights observations by their contribution to identifying variation. If the effect of competition varies across markets, the pooled coefficient is a variance-weighted average of local effects rather than a simple average. This could raise a concern if that variation was disproportionately concentrated in a small number of observations or systematically related to market features. In this subsection, we examine whether either of these concerns are present in our setting.

We first examine whether the distribution of the identifying shock Z_r varies systematically with market characteristics. Figure A.2 plots the distribution of Z_r against the number of nearby applicants. Figure A.3 plots Z_r against population density. The interquartile ranges are similar across bins in both cases, with no indication that identifying variation is concentrated in a narrow set of high-density locations or systematically related to market size.

Table A.12 assesses whether the implicit weighting affects our estimates. Following Borusyak and Hull (2023), we construct weights, which are inversely proportional to the conditional variance of the shock. We then re-estimate both the controlled OLS and IV specifications. The reweighted estimates are similar in magnitude to the baseline and not statistically distinguishable from them, confirming that our results are not driven by a subset of markets with particularly informative lottery environments.

Figure A.2: Variation in Recentered Shock by Number of Applicants in 1km

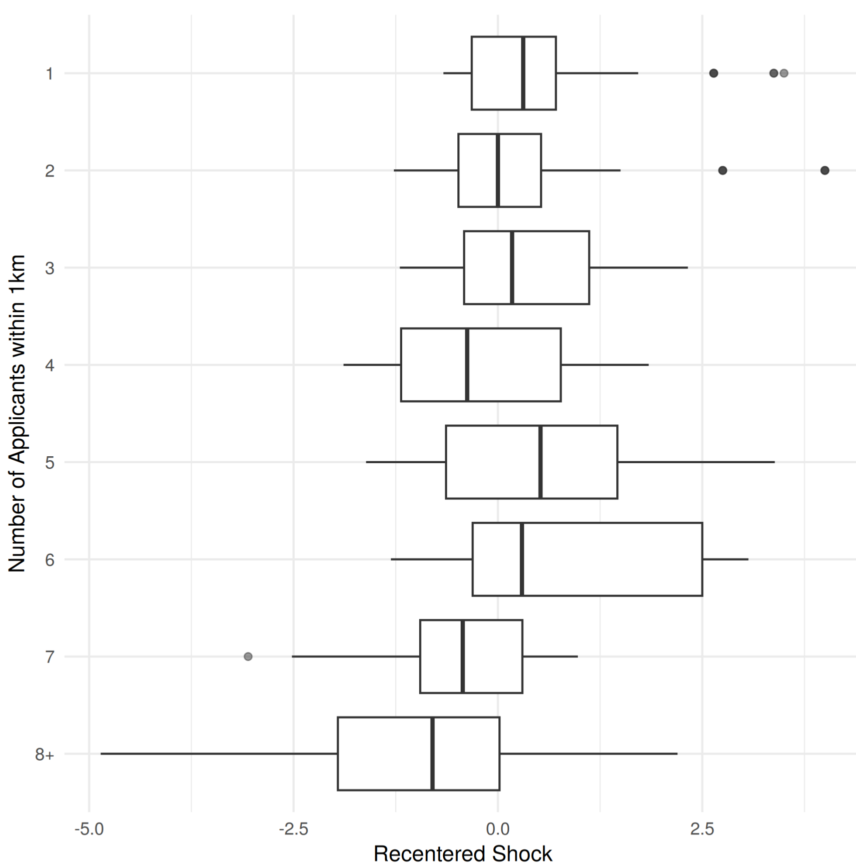


Figure A.3: Variation in Recentered Shock by Population Density

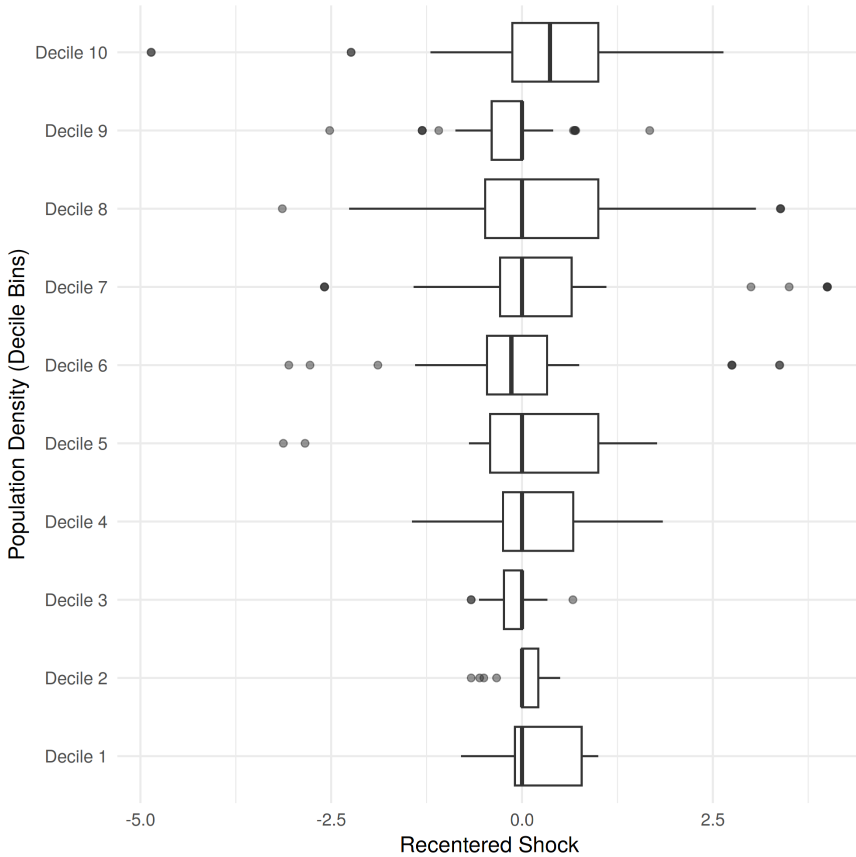


Table A.12: Effect of Retail Competition on Markups: Weighted OLS and IV

	Base Controlled OLS	Weighted OLS	Base IV	Weighted IV
# Comp (1km)	–0.298***	–0.346***	–0.358***	–0.455***
	(0.035)	(0.062)	(0.067)	(0.104)
N	189,483	124,741	189,483	124,741
City-product FE	yes	yes	yes	yes
Quarter-product FE	yes	yes	yes	yes

Notes: Observations for which Var(Z_r) is zero or undefined are excluded by construction.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$