Price Dispersion and the Role of Stores[1]

Espen R. Moen
BI Norwegian Business School, NO-0442 Oslo, Norway
espen.r.moen@bi.no

Fredrik Wulfsberg
Oslo Business School, NO-0130 Oslo, Norway
frewul@oslomet.no

Øyvind Aas
Université libre de Bruxelles, 1050 Bruxelles, Belgium
oyvind.nilsen.aas@gmail.com

Abstract

This paper studies price dispersion in the Norwegian retail market for 766 products across 4 297 stores over 60 months. Price dispersion for homogeneous products is significant and persistent, with a coefficient of variation of 37 percent for the median product. Price dispersion differs between product categories and over time. Store heterogeneity accounts for 30 percent of the observed variation in prices for the median product-month and for around 50 percent for the sample as a whole. Price dispersion is still prevalent after correcting for store heterogeneity.

Keywords: Price dispersion, retail prices, store heterogeneity.
JEL: D2, D4, E3.

[1]* We are grateful for comments from Birthe Larsen, Tore Nilssen, Magnus Söderberg, the referees, and by seminar participants at the 2014 Annual Meeting of the Norwegian Association of Economists, the 8th Nordic Summer Symposium in Macroeconomics, the 9th Nordic Workshop in Industrial Organization at the University of Oslo, the 29th EEA-ESEM meeting at Toulouse School of Economics, Norwegian School of Economics, BI Norwegian Business School, Statistics Norway and Oslo Business School. Øyvind gratefully acknowledges the financial support of the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013 / ERC grant agreement no. 339950).

It is common knowledge that the price for a particular product or service may vary substantially between stores or outlets. One explanation for price dispersion is that stores are different. Stores can be heterogeneous in a multitude of ways such as location, opening hours, parking facilities (see e.g. Dixit and Stiglitz (1977) and Weitzman (1994)), loyalty programs (Basso et al., 2009) and warranties (Grossman, 1981). In addition, idiosyncratic shocks or unexpected fluctuations in demand (Mackowiak and Wiederholt, 2009) may also yield price dispersion. Furthermore, store characteristics are often an intrinsic component of a purchase (product differentiation). For example, buying a lukewarm Coca-Cola in a supermarket in the middle of the day is different from buying a cold one from a convenience store or a petrol station in the middle of the night. Or, eating the same meal at two different restaurants may be perceived as very different depending on characteristics for each restaurant. Store characteristics may thus reflect different mark-ups and costs, and result in price dispersion.

In this paper we identify the contribution of store characteristics to price dispersion exploring monthly price observations from a wider set of products categories than previous studies. But first we establish six stylized facts of price dispersion: (1) there is significant and persistent price dispersion in retail prices in Norway, the median standard deviation is 33 percent of the mean price. (2) The dispersion of prices varies between products and over time as indicated by the range between the first and third quartiles of the standard deviation between 19 percent and 50 percent. There is less price dispersion for non-durable and durable products than for semi-durable products and services. (3) There is little variation in price dispersion between regions. (4) 84 percent of the overall variation in the standard deviation is between products while 16 percent is due to time variation. (5) The dispersion in prices increased from around 25 percent in the start of the sample to almost 40 percent in end of the sample. (6) While the pooled distribution of normalized prices is unimodal, product specific distribution of normalized prices are often bimodal.

Second, we identify a fixed store component of prices by observing prices of multiple products in the same store over time. Using intuitive non-parametric methods we find that the store component accounts for about 30 percent of the price dispersion *for the median product.* To identify the store component for the sample as a whole we also employ a novel parametric method, which shows that store effect accounts for 50–60 percent of the dispersion in prices. As further evidence of the importance of store heterogeneity, we find that the ranking of stores within the price distributions is highly persistent over time.

Kaplan and Menzio (2015) use scanner data from 1.4 million grocery products across different geographical areas in the United States. They find that the quarterly average standard deviation in prices is between 19 and 36 percent depending on the aggregation level,[2] and that store heterogeneity account for 10 percent of the dispersion in prices. Exploring a subset of the data and a different method, Kaplan, Menzio, Rudanko, and Trachter (2016) find that 15.5 percent of the variance in prices is due to store heterogeneity.[3] In a similar study using French grocery data from many stores belonging to a few supermarket chains Berardi, Sevestre, and Thébault (2017) find that the average price dispersion is 7 percent. They find that a permanent component dominates price dispersion between stores due to centralized price setting. Lach (2002) studies price dispersion for only four products[4] in Israel. He finds that store characteristics account for between 47–90 percent of the variation in prices. Wildenbeest (2011) investigate price dispersion of a basket of grocery items from four retailers in the United Kingdom. He finds that store heterogeneity explains around 61 percent of variation in prices and attributes the rest to search frictions.

We contribute to this literature by covering a larger variety of products from multiple stores. We include not only food products, but also products from all 10 COICOP categories such as consumer electronics, cars, petrol, apparel, restaurants, transport, and

[2] Products are defined by universal product code (UPC).

[3] They also decompose the variation in prices into transitory and persistent parts. The persistent component of the store-product variation in prices (which they label "relative price dispersion") constitutes 30.3 percent of the variation in prices.

[4] These are refrigerator, chicken, flour, and coffee.

other services. This allows us to provide more detailed insight into price dispersion than previous studies. While the above cited studies report measures of price dispersion for the overall or pooled price distribution, we report price dispersion at the product level and the extent of variation across products. Also, we argue that the store effect may represent information about the price structure in the market.

I The Role of Stores

Households may choose from a tremendously large set of products of different brands and qualities at different prices from different stores. Considering what products to buy where is a huge task for consumers, and it is almost impossible in practice to gather and process all available information.

Suppose all variation in prices can be attributed to the store effect. This means that the ratio of the prices of any goods in two stores A and B are equal to the ratio of the average price in the two stores. Hence knowing the ratio of the average in two stores then reveals the ratio of the prices of any good in the two stores, and hence also the ratio of the prices of any baskets of goods between the two stores. In effect, a multi-good search problem simplifies to a single-good search problem. If the store effect is substantial but does not explain all the price variation, the ratio of the average price between two stores give some but not full information about the price ratios of individual goods between the two stores.

In many cases it may be relatively easy for consumers to collect information about average price level in a store. Surveys for instance, give information about average prices. The more is explained by the store effect, the easier is it for the customers to be informed about relative price levels. Hence measuring the store effect gives information about how easy it is to get information for customers, which again is a stepping stone for understanding the working of the price system as a whole. The store effect may thus be an indication of the information structure in the market, and a high store effect makes it more likely that the consumers are well informed and that competition works well.

There are several mechanisms that can explain why store effect exists. Quality differences associated with location and opening hours may yield differences in mark-ups or marginal costs between stores. Sellers may have some local market power, and the elasticity of demand facing the sellers and hence also the optimal mark-up may vary between stores. In an environment with imperfectly informed customers and search as in Burdett and Judd (1983), identical stores may charge different average prices. The sellers that charge high average prices have high mark-ups but low sales, while the reverse is true for sellers that offer low average prices, making the expected profit equal. Finally, store effects may be a result of product aggregation. As some products in our data may represent different brands and qualities across stores, this may yield a store effect. In our empirical analyses below we do not distinguish between the different sources of the store effect.

The next question is why store effects do not explain all dispersion in prices? First, it may be that the marginal cost of selling different products vary. A store with a huge freezer may have lower marginal costs of storing and selling frozen products than a store with a small freezer. Related, firms may specialize in selling some products, and set higher mark-ups on these products. They may specialize in expertise to help consumers choose between products of different qualities and properties while selling more standardized products at discounted prices. A store selling cross country skis may specialize in selecting the optimal pair of skis (with the appropriate span and stiffness) for each individual customer, and thus charge a higher price on skis while charging the same price as a general store for ski wax.

Within-store price dispersion may also be explained by price discrimination. Kaplan, Menzio, Rudanko, and Trachter (2016) analyze a two-goods variant of the model of Burdett and Judd (1983). Customers differ in the cost of shopping time. By selling the two goods at different prices, the store may be able to sell the bundle at a high average price to customers with a high cost of shopping, and in addition sell the low-priced good to customers with a low cost of shopping. This leads to within-store dispersion.

Lal and Matutes (1989) study a very similar mechanism. They consider a duopoly framework in which both sellers sell the same two goods, while customers differ in their transportation cost (the cost of going to the both stores). In one (out of several) equilibrium, sellers sell one of the goods at a discount, which one differs between the stores. The customers with a high transportation cost buy both products in the same store, while those with a low transportation cost buy one good at each store.

Sellers with market power that serve a given customer pool may also engage in temporal price discrimination. This is particularly relevant for durable goods. Most of the time prices are set high, and the sellers sell to high-valuation customers only. However, from time to time the seller sets prices low (sales) in order to serve segments with lower willingness to pay, see Conlisk, Gerstner, and Sobel (1984).

Also, incomplete information may yield a search equilibrium where prices differ randomly between stores cf. Burdett and Judd (1983) and Moen (1999).[5] Lastly, Kaplan and Menzio (2015) shows that price discrimination also may yield price dispersion beyond store effects.

II Data

We use monthly micro data collected for the consumer price index (CPI), see Statistics Norway website[6] and Statistics Norway (2001). The data covers monthly price observations of 766 different products and services from 4,297 stores from January 2000 to December 2004. In total, our sample consists of 2,774,494 price observations. The median number of observed prices in a store (in one month) is 46 with and IQR between 19 an 187.[7] Appendix A reports some further descriptives on the size of the data set.

The products represent all COICOP[8] divisions such as food, apparel, furnishing, trans-

[5] Moen (1997) shows that price dispersion also may emerge with directed search.

[6] `http://www.ssb.no/en/priser-og-prisindekser/statistikker/kpi/`

[7] The observed prices are a sample of all the goods sold by a store. Our underlying assumption is that the price of the basket of sampled goods is representative for the overall price level in that store. This seems reasonable given that the purpose of collecting price information is to measure CPI.

[8] COICOP is an acronym for Classification of Individual Consumption According to Purpose, which is a nomenclature developed by the United Nations Statistics Division to classify and analyze individual

port, services, recreation, electronics, and fuels to name a few. Compared to Kaplan and Menzio (2015) and Kaplan, Menzio, Rudanko, and Trachter (2016) who analyze grocery prices (mostly products categorized by COICOP division 1) while our data includes a larger variety of products covering more than 70 percent of household expenditures. The panel is unbalanced since some products and stores are replaced each year by Statistics Norway to ensure the representativeness of the consumption basket (Statistics Norway, 2001). The products are defined with varying degree of precision depending on its type. For example, *Coca-Cola, bottle, 0.33 liters* is more precise than *Dress, ladies, simple.* Thus, for some products price dispersion may reflect differences in quality between stores. The products are thus a sample and not the population of the products in each store.

Each store is located in one of 8 regions: (1) Oslo, (2) Bergen, (3) Trondheim, (4) Akershus county, (5) Eastern Norway, (6) Southern and Western Norway, (7) Central Norway, and (8) Northern Norway. The regions are heterogeneous differing in geographical size and population density. Regions (1)–(3) are main cities, region (4) is a county just east of Oslo, while regions (5)–(8) are larger entities, see the map in Figure A1 in the appendix.

Some prices may change more often than each month so there may be some dispersion in prices these data will miss. However, using the same data, Wulfsberg (2016, Table B2) reports that the mean duration of a price spell (time between two price changes) varies between 5.8 months for food products (COICOP division 1) and 16.0 months for transport services (COICOP division 7).

III Stylized Facts of Price Dispersion

In this section, we present different measures of price dispersion in our sample and how dispersion varies between products and over time. We denote P_{ist} as the price observation for product i in store s at month t. First, we construct a price distribution for each

consumption expenditures incurred by households according to their purpose.

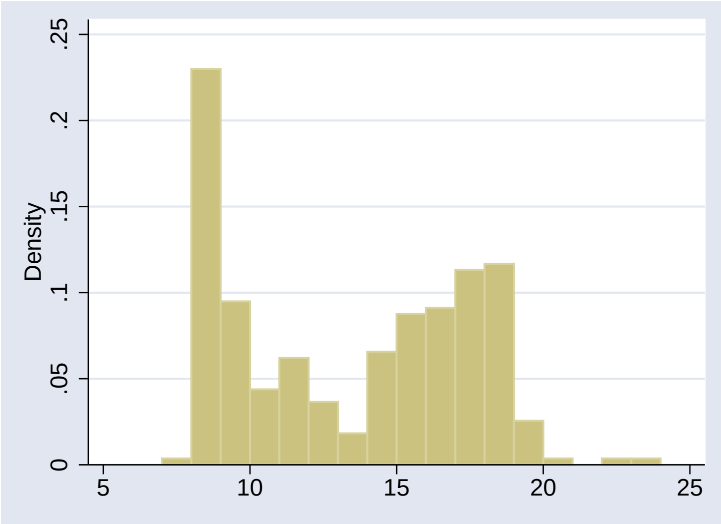

Figure 1: The distribution of prices in NOK of Coca-Cola in January 2000.

product-month sample, in total 40,567 distributions.[9] We drop product-months with fewer than 20 observations (stores) in order to reduce sampling errors. The median number of observations in a product-month distribution is 58 with an inter quartile range of (39–87).

As an example, Figure 1 shows the distribution of observed prices P_{ist} for a bottle of Coca-Cola (in NOK) from 274 stores in January 2000. The lowest price is 7 NOK and the highest price more than three times higher at 24 NOK. The third quartile price is 89 percent higher than the first quartile price, and the standard deviation is 3.70 NOK which is 28 percent relative to the mean price. Interestingly, the price distribution is clearly bimodal. One possible interpretation of this feature is that stores are either "cheap" (like supermarkets) or "expensive" (like convenience stores).

Each product-month distribution of P_{ist} will obviously have different means μ_{it} and variances Σ_{it} which may depend on the the scale (i.e. the mean μ_{it}). In order to compare the dispersion of prices across products and months we normalize all prices P_{ist} with

[9]Unfortunately there are not enough observations to construct price distributions at the product-region-month level. If we were to keep regions as one of the dimensions, we would either need to drop the product or time dimension, i.e. to analyze region-month or region-product distributions. As documented, price dispersion varies a lot between products, but it is also important to control for idiosyncratic changes in the mean product price over time. Furthermore, by analyzing product-time distributions we can directly compare our results with previous studies. Hence, we stick to analyzing product-month distributions in our main analysis. This also allows us to compare our results with previous studies.

respect to the mean price for each product in each month $\bar{P}_{it}$:

$$\tilde{P}_{ist} = P_{ist}/\bar{P}_{it} \tag{1}$$

$\tilde{P}_{ist}$ has thus a mean of 1 and variance σ^2_{it}, which we can compare across products and over time.[10]

As the product-month price distributions may be skewed or multimodal like the distribution for Coca-Cola in Figure 1, we measure price dispersion as the interquartile range (IQR) and the 95/5 percentile range (P95–P5 range) in addition to the variance and standard deviation. Since the distributions are normalized, the standard deviation, the IQR and the P95–P5 range are in percent of the mean.

The first column of Table 1 reports the median product-month estimates of these measures, and in column 2 we illustrate the variation in each of these measures (across products and over time) by the range between the first and third quartile, i.e. the ($Q1$–$Q3$) range. Table 1 shows that price dispersion is significant with a estimated median standard deviation of 32.7 percent (corresponding to a variance of 0.107). The median IQR is 31.9 percent and the median ($P95-P5$) range is 94.5 percent. However, there is a lot of variation in dispersion between the product-month distributions as indicated by a ($Q1-Q3$) range of the standard deviation between 19 percent and 50 percent. Similarly there is a lot of variation in the other measures of price dispersion. The ($Q1-Q3$) range for IQR is between 16.6 percent and 55.8 percent, and between 54.8 percent and 149.0 percent for the ($P95$–$P5$) range. This variation may reflect that some products are precisely defined while others are aggregates of products that are close substitutes. Figure 2 shows a histogram of the 40,567 standard deviations illustrating the variation in price dispersion. The distribution is skewed right, which is why we report on the median.[11]

Grouping the products by COICOP categories shows systematic differences in disper-

[10]We exclude 4010 outliers i.e. 0.14 percent of the observations, with a relative price greater than 5 or less than 0.05 and then renormalize. The outliers represents all regions and COICIOP divisions nonsystematically.

[11]The mean variance and standard deviation are 0.180 and 36.3 percent.

Table 1: Descriptive statistics for relative price dispersion across products and over time.

Dispersion measure	Median	$(Q1-Q3)$ range
Standard deviation	0.327	$(0.193-0.504)$
IQR	0.319	$(0.166-0.562)$
$(P95-P5)$ range	0.945	$(0.548-1.480)$

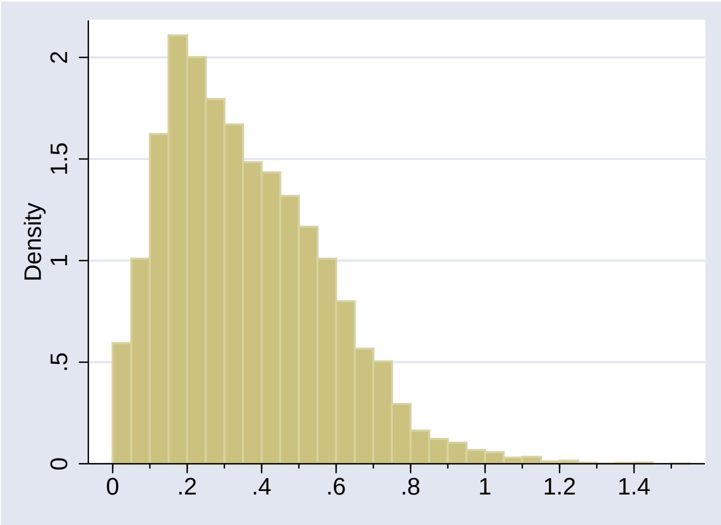

Figure 2: The histogram of the standard deviation for all product-month distributions truncated at 2.

sion between types of products. Table 2 shows the median standard deviation and IQR for each COICOP division in the top panel and the degree of durability in the bottom panel. The groups are ranked according to their standard deviation. *Clothing and footwear* products have the highest price dispersion with a median standard deviation of 55.1 percent. The least dispersed categories are *Health* products and *Alcoholic beverages and tobacco* with a median standard deviation of less than 10 percent. In particular, the dispersion we find in normalized prices for food products in Norway is 25.3 percent (measured by the median standard deviation), which is similar to what Kaplan and Menzio (2015) and Kaplan, Menzio, Rudanko, and Trachter (2016) find for grocery products in the United States (19–36 percent).

The COICOP system also classifies products as durables, semi-durables, nondurables and services, see Table A2 in the appendix. For example, clothing and footwear pro-

Table 2: Median dispersion in relative prices across COICOP *categories. Ranked by median standard deviation.*

		Standard deviation	
COICOP division	N	Median	$(Q1 - Q3)$ range
3 Clothing and footwear	102	0.551	(0.425 – 0.666)
5 Furnishings, household equip	120	0.424	(0.254 – 0.584)
11 Restaurants and hotels	38	0.408	(0.307 – 0.488)
8 Communication	7	0.401	(0.343 – 0.523)
7 Transport	37	0.351	(0.161 – 0.515)
9 Recreation and culture	81	0.325	(0.199 – 0.488)
12 Miscellaneous goods, services	60	0.325	(0.219 – 0.536)
1 Food and non-alcoholic bevs	255	0.253	(0.170 – 0.367)
4 Housing, water, electricity, gas and other fuels	15	0.250	(0.134 – 0.385)
2 Alcoholic beverages, tobacco and narcotics	12	0.099	(0.068 – 0.179)
6 Health	39	0.089	(0.035 – 0.223)
Semi-durables	184	0.508	(0.384 – 0.644)
Services	72	0.388	(0.248 – 0.497)
Durables	86	0.372	(0.213 – 0.553)
Non-durables	424	0.243	(0.154 – 0.376)

Note: N is number of products in each category.

ducts are classified as semi-durables. The bottom part of Table 2 shows the median product-month standard deviation within these categories. The standard deviation for semi-durable products is about twice as high as for nondurable products (50.8 percent vs. 24.3 percent). The dispersion of services and durables are in between at around 38 percent. The right column of Table 2 shows that there is a lot of heterogeneity within each consumption category indicated by the $(Q1 - Q3)$ range of the standard deviation.

Price dispersion is positively correlated with the price level with a correlation coefficient of 0.20 between the standard deviation and the log of the mean price. However, the relationship is actually hump shaped, see figure 3. The figure is replicated in colors as figure C2 in the appendix showing that non-durables dominate the low-price segment, du-

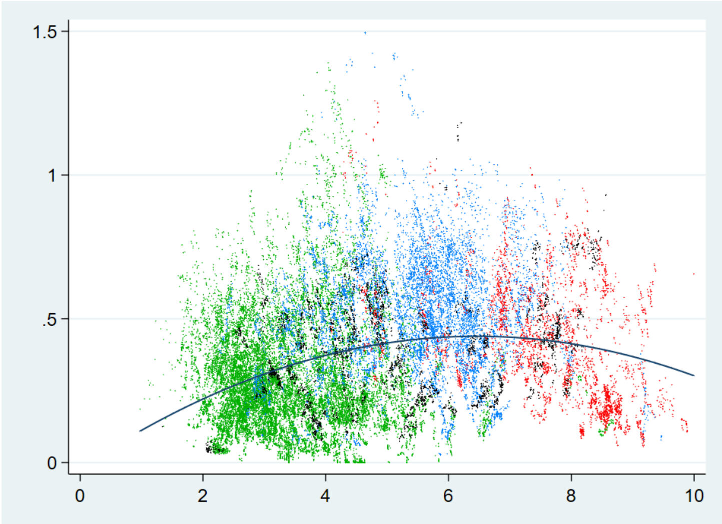

Figure 3: Scatter plot of the standard deviation (verical axis) and log of the mean price (horisontal axis) with a quadratic fitted line. Non-durables are green, semi-durables are blue, durables are red, and services are black.

rables dominate the high-price segment, while semi-durables are in between. The average price of services are represented over the whole price range. These systematic differences shed light over the difference in price dispersion across COICOP divisions in Table 2. In particular COICOP divisions 3 *Clothing and footwear* and 5 *Furnishings, household equipment* are dominated by semi-durable products.

Figure 4 shows that there is a clear upward trend in the three quartiles (Q1, median and Q3) of the standard deviation indicating that dispersion increased over time. The median standard deviation increased from around 25 percent in the beginning of the sample to almost 40 percent in the end. This finding is consistent with Wulfsberg (2016) who reports that the mean size of nonzero price changes increased over the same period.

How much of the variation in price dispersion reported in Table 1 is accounted for by this trend? Decomposing the variation in the standard deviation σ_{it} into cross sectional variation between products ($\bar{\sigma}_i$) and time variation within products ($\sigma_{it} - \bar{\sigma}_i$) yields that the cross sectional variation accounts for 84 percent while time variation only accounts for 16 percent of the overall variation in σ_{it}.

What can explain this trend? In a Ss-type menu cost model the size of price adjustments increase with inflation. Hence, we would expect more price dispersion when

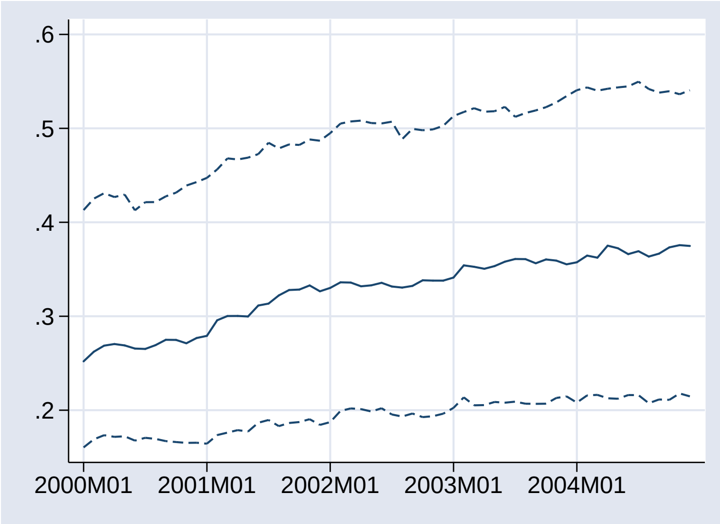

Figure 4: The first, second (median) and third quartiles of the standard deviations over time.

inflation is high. However, during this period the 12 month annual inflation rate has varied between 4 percent and −2 percent with a slightly negative trend (if any). The trend increase in price dispersion may be related to a changing composition of products change over time. To investigate this possibility, we have estimated price dispersion for the subset of products that are observed over the whole period. Table E1 in the appendix shows that the estimates of price dispersion for the balanced panel is hardly changed, and figure E1 shows that the trend prevails which indicates that the trend is not explained by changes to the composition of products.

Since the distributions are constructed at the product-month level we cannot report estimates of price dispersion across regions based on these distributions. However, to get an idea on the regional variation on price dispersion we have also constructed normalized price distributions for product and regions (omitting the time dimension), i.e. $\tilde{P}_{isr}$ where subscript r is an index for region. Table D1 in the appendix shows that price dispersion does not vary much between regions. Surprisingly, given the huge geographical size, Northern Norway is the region with the lowest price dispersion of 27.8 percent and Southern and Western Norway has the highest price dispersion of 33.1 percent. Also surprisingly perhaps, is that there is no less price dispersion in the three cities than in the geographically larger regions. Thus regions does not seem to play an important role in price dispersion.

We noted above that a possible explanation for the bimodality of the price distribution for Coca Cola in Figure 1 is that the product is sold by stores with different characteristics such as "cheap" stores (supermarkets) and "expensive" stores (convenience stores). Multimodal product-specific distributions of normalized prices seems to be prevalent. Instead of visual inspection of each product we employ the Hartigan dip test (Hartigan and Hartigan, 1985) which rejects unimodality at 1 percent level of significance for as many as 576 (75 percent) of the products.[12] Another explanation for multimodality is that there are regional differences in the price distributions. Figure D1 in the appendix shows, however, that regional distribution of prices for Coca Cola are still strongly bimodal.

While product specific distributions often are multimodal, the pooled distribution of normalized prices is single peaked, kurtotic, and slightly left skewed as seen from Figure B1 in the appendix. The standard deviation of the pooled distribution is 39.9 percent, which is 7.2 percentage points higher than the median product-month standard deviation.

IV The Store Component

We assume that the relative price $\tilde{P}_{ist}$ can be decomposed into the mean (by definition equal to one), a store component v_s and a residual ε_{ist}:

$$\tilde{P}_{ist} = 1 + v_s + \varepsilon_{ist} \tag{2}$$

The store component is assumed to be equal for all products sold in the same store s in all periods. The idea of the store component is illustrated in Figure 5 where the dots represent four observations of the relative price for a Coca Cola sold in one particular store s from December 2000 to April 2001. The distance from the mean relative price (equal to 1) to the dashed line represents the store component, v_s. The conditional mean

[12]Cavallo and Rigobon (2012) uses the dip test to inspect the distribution of price changes.

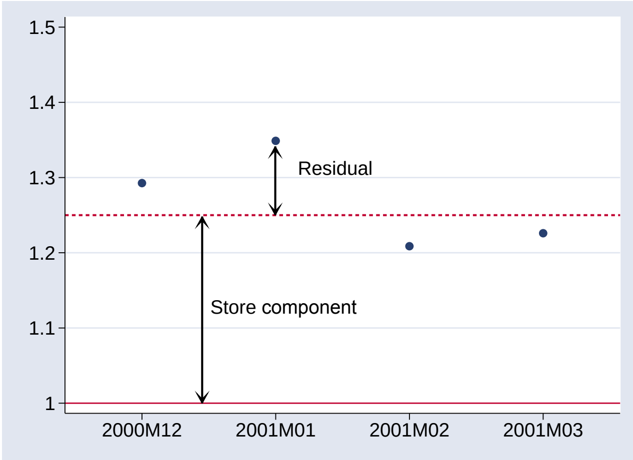

Figure 5: Illustration of the store component. The vertical axis is the relative price and the horizontal axis is the month. The dots are observations from the same store.

of the relative price for all products sold by this particular store, is thus $1 + v_s$. Since the store component is around .25 the store is on average 25 percent more expensive compared to other stores. However, the four price observations vary around the store mean $(1 + v_s)$, represented by the residual ε_{ist}.[13]

An intuitive estimate of the store components v_s is the mean normalized price for all products in all periods sold by the same store:

$$\hat{v}_s = \frac{1}{N_s} \sum_n (\tilde{P}_{ist} - 1) \tag{3}$$

where $n = 1, \ldots, N_s$ is the number of observations from store s over all products and months. Note that the estimation of the store effect is based on a large number of observations. As the median number of products per store is 46 and the median store is observed over 31 months, the median number of observations behind the estimate of the store component is 1426 observations.

A t-test shows that one fifth of the store effects are insignificant at the 1 percent level.[14] The significant (non-zero) store effects are plotted in Figure 6. Their size vary

[13]While the store effect in (2) is fixed over time, Kaplan, Menzio, Rudanko, and Trachter (2016) estimate a time varying store effect by decomposing the error terms further into a transitory and a persistence component. Their results indicate that 95 percent of the sample store effect is persistent.

[14]15 percent of the store effects are insignificant at the 5 percent level.

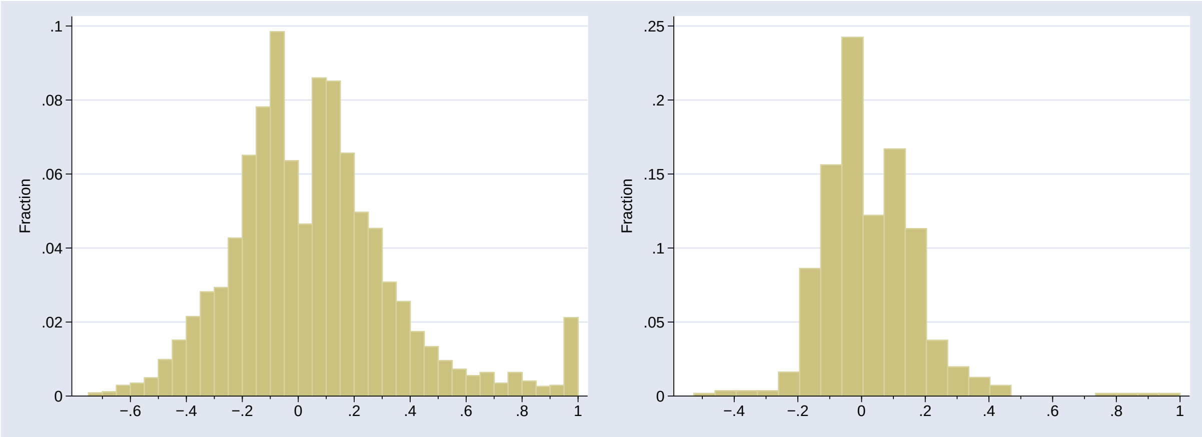

Figure 6: Histograms of the significant store effects (1 percent level significance) for all stores (left) and Coca-Cola stores (right). The histograms are truncated at 1.

between −70 percent and 300 percent. The histogram is clearly bimodal with one mode below zero around −10 percent and one mode above zero around 15 percent. The mean store effect is 28.5 percent for expensive stores and −18.0 percent for cheap stores. (The mean absolute size of the store effects is 23.6 percent.) The histogram to the right is for stores selling Coca-Cola which is also possibly bimodal with modes on each side of zero. In Figure F2 in the appendix the store effects are plotted by COICOP division. Bimodality is also reflected in these histograms with the exception of *9 Recreation and culture* and *12 Miscellaneous goods, services*. We note that the store effects seem particularly strong for *3 Clothing and footwear*.

Interestingly, there is a clear tendency that the more expensive the store is the more variation in (normalized) prices within the store. Figure 7 plots the store effects vs the variation in prices between products sold in the same store. One possible explanation for this is that expensive stores selling specialized products also need to sell standardized products (like our example above of a store selling cross-country skis and ski wax).

The residual component $\hat{\varepsilon}_{ist}$ is computed by subtracting the estimated store effects from each normalized price i.e. $\hat{\varepsilon}_{ist} = \tilde{P}_{ist} - 1 - \hat{v}_s$ following equation (2). The variance of ε_{ist} represents the price dispersion for products sold at equally expensive stores. Table 3 reports the same measures of residual price dispersion as for normalized prices in Table A1. Lach (2002) which focus on the dispersion of residual prices, finds less dispersion (but

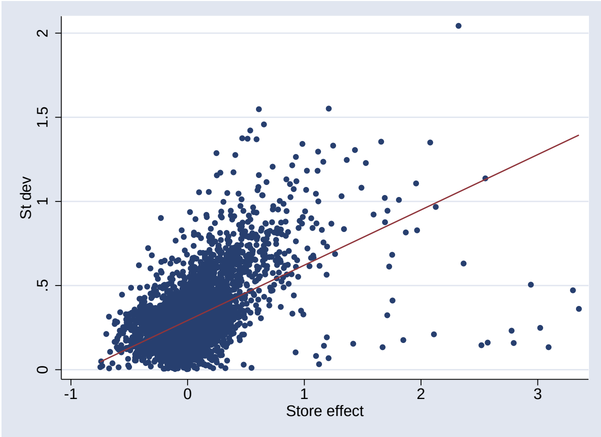

Figure 7: Expensive stores have higher price variation.

Table 3: Descriptive statistics for residual price dispersion across products and over time.

Dispersion measure	Median	$(Q1 - Q3)$ range
Standard deviation	0.274	(0.174 – 0.392)
IQR	0.277	(0.157 – 0.416)
$(P95-P5)$ range	0.806	(0.511 – 1.185)

only for four products). All three dispersion measures of the residual prices are around 85 percent of the corresponding measure for normalized prices. There is thus substantial variation in prices even after controlling for store effects.

V The Importance of Store Heterogeneity

In this section, we explore how much of the variation in prices which we documented in the previous section, can be attributed to store heterogeneity. We interpret the statistical model (2) as an error component model where the store effect v_s and the residual price ε_{ist} are stochastic each with a 0 mean, and a time and product specific variance σ^2_{vit} and $\sigma^2_{\varepsilon it}$. Note in particular that the variance of the store effect may vary between products and over time even if the store effect v_s does not vary between products and over time. The reason for this is that the product range varies between stores over time. To illustrate this point assume that there are in total three products A, B, and C. There are many

Table 4: Estimates of variance components. Share of total variance in parenthesis.

Variance:	Median		$(Q1 - Q3)$ range	
store $\hat{\sigma}^2_{vit}$	0.032	(30)	(0.007 – 0.100)	(19–39)
residual $\hat{\sigma}^2_{eit}$	0.075	(70)	(0.030 – 0.154)	(81–61)
total $\hat{\sigma}^2_{vit} + \hat{\sigma}^2_{eit}$	0.107	(100)	(0.037 – 0.254)	(100)

stores, but each store sells only two products (A,B), (A,C) or (B,C). The store effect for a particular store is equal for both products sold in that store. However, the variance of the store effect for stores selling (A,B) products may be different from the variance of the store effect for stores selling the (A,C) products or the (B,C) products. Hence, the variance of the store effect for product A may differ from product B and product C. More generally, the store effect is drawn from a distribution, which depends on the set of products sold by the store and on the period.

Assuming that v_s and e_{ist} are independent, the variance of $\tilde{P}_{ist}$ is thus decomposed into

$$\sigma^2_{it} = \sigma^2_{vit} + \sigma^2_{\varepsilon it} \tag{4}$$

The ratio of the variance of the store component σ^2_{vit} to the total variance σ^2_{it} measures the importance of store heterogeneity for price dispersion. Our goal is thus to estimate σ^2_{vit} and $\sigma^2_{\varepsilon it}$ and their implied share of the total variance σ^2_{it}.

We first estimate the variance components for each product-month distribution based on the estimates $\hat{v}_s$ and $\hat{\varepsilon}_{ist}$ according to

$$\hat{\sigma}^2_{vit} = \sum_s (\hat{v}_s - \bar{v}_s)^2/(S_{it} - 1) \tag{5}$$

$$\text{and} \qquad \hat{\sigma}^2_{\varepsilon it} = \sum_s (\hat{\varepsilon}_{ist} - \bar{\varepsilon}_{ist})^2/(S_{it} - 1) \tag{6}$$

where $s = 1, \ldots, S_{it}$ is an index for stores selling product i in month t.

In column 1 of Table 4 we report the median variance of the store component and the median residual price variance, which turn out to be 0.032 and 0.075. This yields a total median variance of normalized prices of 0.107 (which is consistent with the estimates in Table 1). Thus the store effect accounts for 30 percent of the total variance of the median product-month. The ratio of the store variance varies obviously between product-month distributions. To illustrate this variation we report the same statistics for the inter quartile range (Q1–Q3 range) in the second column. We see that the store effect accounts for 19–39 percent of the total variance of product-month distributions measured by the Q1–Q3 range.

In Table 5 we report the variance decomposition by COICOP categories using the same approach as in Table 4. The store effect is particularly important for *11 Restaurants and hotels* accounting (49 percent) in addition to *3 Clothing and footwear* (45 percent). Typically for services we would expect variation in the store component to be an important part of the price dispersion. The store effect is least important for *8 Communication* with a ratio of 12 percent to the total variance. For food products (*1 Food and non-alcoholic beverages*) the store effects account for 23 percent of the total variance which is similar to Kaplan and Menzio (2015).

It is likely that stores selling the same product(s) are less heterogeneous than stores in general. For example, food stores are probably less heterogeneous than food stores versus hotels. Hence, for retail prices in general one may expect the variance in the store component to be more important for price dispersion than for the median product-month sample. To investigate this possibility we pool the sample and assume that $\sigma_{it}^2 = \sigma^2$, $\sigma_{vit}^2 = \sigma_v^2$ and $\sigma_{\varepsilon it}^2 = \sigma_\varepsilon^2 \; \forall \; i, t$. In this exercise we estimate σ_v^2 and σ_ε^2 by

$$\hat{\sigma}_v^2 = \sum_s (\hat{v}_s - \bar{v}_s)^2/(S-1) \tag{7}$$

$$\text{and} \qquad \hat{\sigma}_\varepsilon^2 = \sum_n (\hat{\varepsilon}_{ist} - \bar{\varepsilon}_{ist})^2/(N-1) \tag{8}$$

where $s = 1, \ldots, S$ is an index for all stores, and $n = 1, \ldots, N$ is an index for all

Table 5: Variance decomposition by COICOP division. Mean estimates.

COICOP division	σ^2	σ_v^2	σ_v^2/σ^2
11 Restaurants and hotels	0.151	0.074	49%
3 Clothing and footwear	0.333	0.150	45%
12 Miscellaneous goods, services	0.181	0.057	32%
2 Alcoholic beverages, tobacco and narcotics	0.055	0.017	32%
5 Furnishings, household equip	0.222	0.066	30%
4 Housing, water, electricity, gas and other fuels	0.098	0.026	26%
9 Recreation and culture	0.143	0.036	25%
6 Health	0.078	0.018	23%
1 Food and non-alcoholic bevs	0.109	0.025	23%
7 Transport	0.196	0.035	18%
8 Communication	0.189	0.022	12%
Services	0.159	0.060	38%
Semi-durables	0.293	0.106	36%
Durables	0.191	0.067	35%
Non-durables	0.112	0.027	24%

Table 6: Estimates of variance components. Pooled distribution

Variance:	Pooled		ME	
store $\hat{\sigma}_v^2$	0.104	(66)	0.098	(48)
residual $\hat{\sigma}_e^2$	0.053	(34)	0.107	(52)
total $\hat{\sigma}_v^2 + \hat{\sigma}_e^2$	0.157	(100)	0.205	(100)

observations in the sample. As reported in column 1 of Table 6 this yields an estimate of the variance of the store component of 0.104, which is significantly larger than the estimate in Table 4 as expected. This accounts for 66 percent of the pooled total variance leaving 34 percent for the residual variance, indicating a significant larger role for store effects when we look at the whole sample of products.

For robustness we estimate σ_v^2 and σ_e^2 directly using the mixed effects (ME) method (see Rabe-Hesketh and Skrondal, 2012 for details). This method estimates $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$ simultaneously using maximum likelihood without first estimating the store component

v_s, but assuming normality and homoskedasticity, i.e. $v_s \sim N(0, \sigma_v^2)$. This approach yields an estimate of the store component variance σ_v^2 equal to 0.098 which is a share of 48 percent of the total variance, see column 2 of Table 6. Note that the ME estimate of the total variance is larger than the pooled variance because of the normality assumption while the empirical (pooled) distribution is kurtotic (see Figure B1 in the appendix).

The different approaches thus yield an estimated share of the store effects from 30 percent for the median product-month sample to around 50 percent for the pooled sample. Our results thus attribute a somewhat stronger importance to store heterogeneity than Kaplan, Menzio, Rudanko, and Trachter (2016) who attribute 10–36 percent of the observed price dispersion to store heterogeneity. The main reason for this difference is that we analyze prices for a wider product range and probably more heterogeneous stores.

Persistence

Store heterogeneity is an important component of the observed price dispersion, as documented above. In order to investigate the persistence of the store heterogeneity we inspect the ranking of stores within the price distributions over time following Lach (2002). Is a store's ranking in the price distribution persistent as indicated by the estimated store effects?

For each product-month, we partition each price distribution by the three quartiles ($Q1$, median, and $Q3$) and assign each store into one of the four quartile bins $QB1_{it}$, $QB2_{it}$, $QB3_{it}$, and $QB4_{it}$.[15] For how long does a store remain in the same quartile bin? Furthermore, how likely is it that a store which changes its nominal price, jumps from one quartile bin to another or remains in the same part of the relative price distribution? If a store is systematically more expensive, consumers can learn this information and take advantage of price differences. If some consumers are informed about prices while others are not (Varian, 1980) shows that it is optimal for a store to randomize its price.

[15] Random sampling errors will, however, induce statistical errors in our ranking of stores.

Denote the probability for store s of moving from QBk_{it} to QBj_{it+1} as γ_{kji}. We estimate γ_{kji} by the fraction of stores selling product i which moves from quarterly bin k in month t to quarterly bin j in the next month.[16] For each product i the transition probabilities γ_{kji} give us all the elements of the one month transition probability matrix. Table 7 reports the one month transition probability matrix for the median product.[17] Each row represents the probability of either staying in the same bin or moving to another bin.[18] For example, the median probability of moving from the first quartile bin $QB1$ to the second quartile bin $QB2$ in the next month is $\gamma_{12} = 7.9$ percent.

We see that a store is most likely to stay in the same quartile bin in the next month since the probabilities along the diagonal are the largest varying between 83–93 percent. Given that a store moves from one quartile bin to another, it is most likely to move to an adjacent bin. The closest elements to the diagonal vary between 6.2–8.0 percent. The probability of jumping two quartile bins e.g. from $QB3$ to $QB1$ ranges between 0.8–1.5 percent. A store is least likely to move for one tail to the other. The median probability of moving from the lowest to the highest bin is 0.6 percent while the probability of the reverse transition is 0.3 percent. The matrix is quite symmetric, but the upper elements are somewhat larger than the corresponding lower elements. This indicates that the likelihood of moving down from for example the third quartile bin to the first quartile bin γ_{31} is smaller than moving up from the first quartile bin to the third quartile bin γ_{13}. The transition probability matrix varies across products, as indicated by the standard deviations.

We find the same pattern when we estimate the 12 month median transition probability matrix, see Table 8.[19] Even 12 months ahead a store is most likely to remain in the same quartile bin than move to any other bin. The median probability of being in the

[16] This estimate corresponds to the predicted probabilities of a probit model with the initial quaterly bin as regressor.

[17] See Table G1 in the appendix for the mean probabilities.

[18] Note that the rows does not sum to one since each element is the median value of each cell. However, the rows sum to one for each individual product.

[19] See Table G2 in the appendix for the mean probabilities. Table G3 in the appendix reports the 6 month transition probability matrices for the median, the mean, and the standard deviation.

Table 7: 1 month transition probability matrix, normalized prices $\tilde{P}_{ist}$. Median estimates with standard errors in parenthesis. All observations.

	Destination bin			
Origin bin	$QB1_{t+1}$	$QB2_{t+1}$	$QB3_{t+1}$	$QB4_{t+1}$
$QB1_t$	0.885	0.079	0.015	0.006
	(0.096)	(0.061)	(0.030)	(0.053)
$QB2_t$	0.071	0.830	0.076	0.011
	(0.056)	(0.145)	(0.062)	(0.084)
$QB3_t$	0.010	0.080	0.831	0.069
	(0.023)	(0.059)	(0.152)	(0.109)
$QB4_t$	0.003	0.008	0.062	0.925
	(0.010)	(0.017)	(0.039)	(0.058)

Note: The rows does not sum to one since each element is the median value. However, the rows sum to one for each individual product.

Table 8: 12 months transition probability matrix, normalized prices $\tilde{P}_{ist}$. Median estimates with standard errors in parenthesis. All observations.

	Destination bin			
Origin bin	$QB1_{t+12}$	$QB2_{t+12}$	$QB3_{t+12}$	$QB4_{t+12}$
$QB1_t$	0.632	0.216	0.065	0.031
	(0.194)	(0.125)	(0.092)	(0.124)
$QB2_t$	0.174	0.506	0.206	0.056
	(0.098)	(0.186)	(0.134)	(0.109)
$QB3_t$	0.046	0.200	0.521	0.202
	(0.056)	(0.099)	(0.178)	(0.158)
$QB4_t$	0.019	0.047	0.183	0.734
	(0.036)	(0.068)	(0.082)	(0.129)

Note: See Table 7

same quartile bin in 12 months varies between 51–73 percent compared to 83–93 percent for the 1 month ahead estimates in Table 7.

A change in a store's ranking within the relative price distribution can happen as a result of not only changing its own price, but also if other stores have changed their price. It is interesting to know the transition probabilities conditional on the store changing its own nominal price. Table 9 reports the conditional one month transition probability

Table 9: 1 month transition probability matrix conditional on nominal price changes, normalized prices $\tilde{P}_{ist}$. Median estimates with standard errors in paranthesis.

	Destination bin			
Origin bin	$QB1_{t+1}$	$QB2_{t+1}$	$QB3_{t+1}$	$QB4_{t+1}$
$QB1_t$	0.722	0.162	0.056	0.025
	(0.156)	(0.086)	(0.070)	(0.116)
$QB2_t$	0.140	0.595	0.171	0.045
	(0.081)	(0.166)	(0.104)	(0.114)
$QB3_t$	0.043	0.158	0.602	0.166
	(0.056)	(0.075)	(0.164)	(0.127)
$QB4_t$	0.017	0.038	0.140	0.788
	(0.035)	(0.044)	(0.063)	(0.096)

Note: See Table 7

matrix. Still, the largest probabilities are found on the diagonal ranging from 60 percent to 79 percent. If a store does change its ranking following a nominal price change, it is most likely to move to an adjacent quartile, with probabilities ranging between 14.0–17.1 percent. These probabilities are roughly double compared to the corresponding unconditional probabilities. The probability of jumping two quartiles, for example from $QB1$ to $QB3$, ranges now between 3.8–5.6 percent, increasing with a factor of 4 compared to the unconditional probabilities. Finally, the median probability of moving between the tail bins are 1.7 and 2.5 percent. The standard deviations are larger than in the unconditional estimation, so there is more variation in the transition probability matrices when we condition on a nominal price change.

Our results suggest there are persistent patterns in the ranking of stores within a distribution consistent with the finding that fixed store effects is an important component of variation in prices. Knowing the ranking of stores from a previous period may imply significant search cost savings for consumers since the previous ranking is a fair bet for the current ranking.

Fixed store effects are likely to be related to the persistence of relative prices. It is thus possible for consumers to learn what stores are cheaper on average. But how is the

Table 10: 1 month transition probability matrix, residual prices $\hat{\varepsilon}_{ist}$. Median estimates and standard errors in parenthesis. All observations.

	Destination bin			
Origin bin	$QB1_{t+1}$	$QB2_{t+1}$	$QB3_{t+1}$	$QB4_{t+1}$
$QB1_t$	0.907	0.069	0.011	0.007
	(0.076)	(0.043)	(0.020)	(0.013)
$QB2_t$	0.069	0.846	0.074	0.011
	(0.041)	(0.093)	(0.043)	(0.018)
$QB3_t$	0.011	0.077	0.844	0.064
	(0.021)	(0.044)	(0.093)	(0.039)
$QB4_t$	0.006	0.010	0.064	0.917
	(0.013)	(0.018)	(0.035)	(0.036)

Note: The rows does not sum to one since each element is the median value. However. the rows sum to one for each individual product.

Table 11: 12 months transition probability matrix, residual prices $\hat{\varepsilon}_{ist}$. Median estimates and standard errors in parenthesis. All observations.

	Destination bin			
Origin bin	$QB1_{t+12}$	$QB2_{t+12}$	$QB3_{t+12}$	$QB4_{t+12}$
$QB1_t$	0.676	0.199	0.064	0.043
	(0.147)	(0.083)	(0.062)	(0.048)
$QB2_t$	0.185	0.529	0.218	0.062
	(0.079)	(0.132)	(0.086)	(0.057)
$QB3_t$	0.062	0.204	0.517	0.191
	(0.060)	(0.087)	(0.136)	(0.092)
$QB4_t$	0.039	0.058	0.183	0.697
	(0.052)	(0.005)	(0.075)	(0.133)

Note: See Table 10

relative price mobility of equally expensive stores? To answer this question we control for the store effects and estimate transition probability matrices for the residual prices $\hat{\varepsilon}_{ist}$? Tables 10–12 report the unconditional transition probability matrices for the residual prices for 1 and 12 months, and 1 month conditional on a nominal price change.[20]

Comparing the diagonal elements of the residual price transition probability matrix

[20]See Table H1 and H2 in the appendix for the mean probabilities. Table H3 reports the 6 month transition probability matrices for the median, the mean, and the standard deviation.

Table 12: 1 month transition probability matrix conditional on nominal price changes, residual prices $\hat{\varepsilon}_{ist}$. Median estimates with standard errors in paranthesis.

	Destination bin			
Origin bin	$QB1_{t+1}$	$QB2_{t+1}$	$QB3_{t+1}$	$QB4_{t+1}$
$QB1_t$	0.769	0.136	0.044	0.028
	(0.129)	(0.072)	(0.051)	(0.040)
$QB2_t$	0.130	0.661	0.146	0.043
	(0.072)	(0.133)	(0.078)	(0.048)
$QB3_t$	0.037	0.135	0.669	0.137
	(0.051)	(0.076)	(0.130)	(0.078)
$QB4_t$	0.024	0.035	0.121	0.809
	(0.036)	(0.043)	(0.068)	(0.104)

Note: See Table 10

in Tables 10 and 11 with the corresponding probabilities for relative prices in Tables 7 and 8, we see that they are even higher for $QB1$ and $QB2$, but a little lower for $QB3$ and $QB4$. Relative prices for equally expensive stores are thus still very persistent. Lach (2002) finds more flexibility for residual prices than our results. However, he is analyzing only four products and with only one product for each store which may lead to a biased store effect is prices are not perfectly correlated within each store.

We also measure the duration of being in a particular bin for stores. Figure 8 presents box-plots (across products) showing the fraction of different spells within each quartile bin. Most spells are typically between 1 to 3 months within either of the quartile bins. But there is also a huge fraction of products where stores remain in the same quartile bin for 12 months or more in particular the lower quartile bin $QB1$ and top bin $QB4$.

The relationship between the ranking spells and the transition probability matrix is the following. The conditional probability of changing to a different quartile is the sum of the off-diagonal elements in the transition probability matrix. Taking the average across the four quartiles, we get the probability of changing a quartile one month ahead. This probability is equal to the probability of observing a one-month spell (Lach, 2002). These probabilities are very similar for each individual product in our estimations.

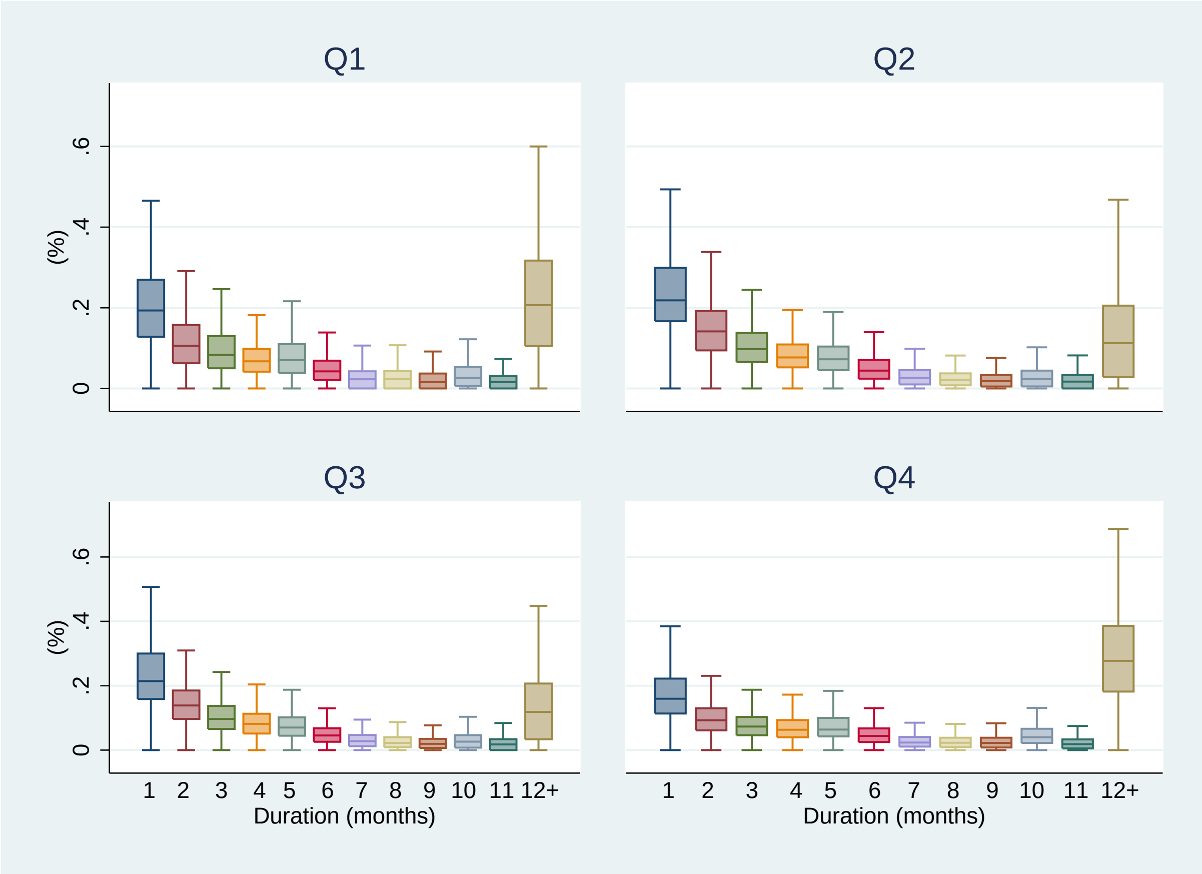

Figure 8: Box-plot of the monthly durations across the four quartiles. Note: See the appendix for the table with the data used to make the box plots. The table consists of mean, median and standard deviations.

Based on the ranking spells and the transition probability matrix, stores in our sample are persistently cheap or expensive. This result, combined with the result from the variance decomposition indicate that store heterogeneity is an important factor for price dispersion.

VI Conclusion

We document empirical facts of price dispersion for a wider range of retail products and services than in earlier studies. The standard deviation for the median product is 33 percent. Dispersion varies between products and months, indicated by the inter quartile range of the standard deviation from 19 percent to 50 percent. Prices appear more dispersed for clothing and footwear and other semi-durable goods than for other products. Furthermore, price dispersion increased over time illustrated by an increase in the standard deviation for the median product from 25 percent to 40 percent over the sample period.

Our results suggest that store heterogeneity is an important component in price dispersion. By decomposing the variance in relative prices into a fixed store component and a idiosyncratic term, we find that 30 percent of the observed variance in relative prices for the median product-month can be account for by store heterogeneity. For the sample as a whole store heterogeneity accounts for 50 percent of the variance in relative prices, which is a larger share than reported in previous studies.

The distribution of the store components bimodal with a long right tail. The mean store effect for cheap stores is –18.0 percent while for expensive stores it is 28.5 percent.

The consistency of a stores ranking within a distribution indicate that most stores are likely to be in the same part of the distribution one month and even 12 months ahead.

From a consumer point of view, it is possible to learn what stores are cheap from searching for prices.

References

Basso, L.J., Clements, M.T., and T.W. Ross (2009). Moral hazard and customer loyalty programs. *American Economic Journal: Microeconomics*, 1(1), 101–123.

Berardi N., Sevestre P., and J. Thébault (2017). The Determinants of Consumer Price Dispersion: Evidence from French Supermarkets. In: Matyas L. (eds) The Econometrics of Multi-dimensional Panels. Advanced Studies in Theoretical and Applied Econometrics, vol 50. Springer, Cham.

Burdett, K. and K. Judd (1983). Equilibrium price dispersion. *Econometrica*, 51(4), 955–969.

Cavallo, A. and T. Rigobon (2012). The distribution of the size of price changes. *Working paper, MIT Sloan.*

Conlisk, J., E. Gerstner, and J. Sobel (1984). Cyclic Pricing by a Durable Goods Monopolist. *Quarterly Journal of Economics*, 99(3), 489–505.

Dixit, A.K. and J.E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, (pp. 297–308).

Grossman, S.J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of law and economics*, 24, 461–483.

Hartigan, J. and P. Hartigan (1985). The dip test of unimodality. *The Annals of Statistics*, 13, 70–84.

Kaplan, G. and G. Menzio (2015). The morphology of price dispersion. *International Economic Review*, 56(4), 1–41.

Kaplan, G., Menzio, G., Rudanko, L., and N. Trachter (2016). Relative price dispersion: Evidence and theory. *NBER working paper series.*

Lach, S. (2002). Existence and persistence of price dispersion: An empirical analysis. *The Review of Economics and Statistics*, 84(3), 433–444.

Lal, R. and C. Matutes (1989). Price Competition in Multimarket Duopolies.. *The RAND Journal of Economics*, 20(4), 516–37.

Mackowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review*, 99, 769–803.

Moen, E.R. (1997). Competitive search equilibrium. *Journal of Political Economy*, 105(2), 385–411.

Moen, E.R. (1999). Education, ranking, and competition for jobs. *Journal of Labor Economics*, 17(4), 694–723.

Rabe-Hesketh, S. and A. Skrondal (2012). *Multilevel and longitudinal modmodel using stata. Volume I: Continuous Responses.* Stata Press.

Statistics Norway (2001). Consumer price index 1995–2000. *Official Statistics of Norway.*

Varian, H. (1980). A model of sales. *American Economic Review*, 70(4), 651–659.

Weitzman, M.L. (1994). Monopolistic competition with endogenous specialization. *The Review of Economic Studies*, 61(1), 45–56.

Wildenbeest, M. (2011). An empirical model of search with vertically differentiated products. *The RAND Journal of Economics*, 42(4), 729–757.

Wulfsberg, F. (2016). Inflation and Price Adjustments: Micro Evidence from Norwegian Consumer Prices 1975–2004. *American Economic Journal: Macroeconomics*, 8(3), 175–94.

Appendix to Price Dispersion and the Role of Stores

by Espen R. Moen, Fredrik Wulfsberg, and Øyvind N. Aas

A Data dimensions

Table A1 presents descriptive statistics for the dimensions of our sample. About 2/3 of the products are observed over the whole period (60 months). While most products are observed over the entire period, no store is observed more than 47 months, with a median of 31 months. There are 108 364 combinations of product and stores in the sample. See Wulfsberg (2016) for further descriptions of the data.

Table A1: Descriptive statistics.

	Median	$(Q1 - Q3)$ range
Observations per product	3,080	$(1,980 - 4,751)$
Number of months per product	60	$(53 - 60)$
Number of months per store	31	$(19 - 47)$
Number of stores per product-month	58	$(39 - 87)$
Number of products per store	46	$(19 - 187)$

Note: Q1 and Q3 are the first and third quartiles.

Figure A1: Observations are connected to one of 8 regions:
(1) Oslo (black spot to the south),
(2) Bergen (black spot to the west),
(3) Trondheim (black spot in Central Norway),
(4) Akershus county (purple county),
(5) Eastern Norway (yellow counties),
(6) Southern and Western Norway (blue counties),
(7) Central Norway (green counties), and
(8) Northen Norway (orange counties).

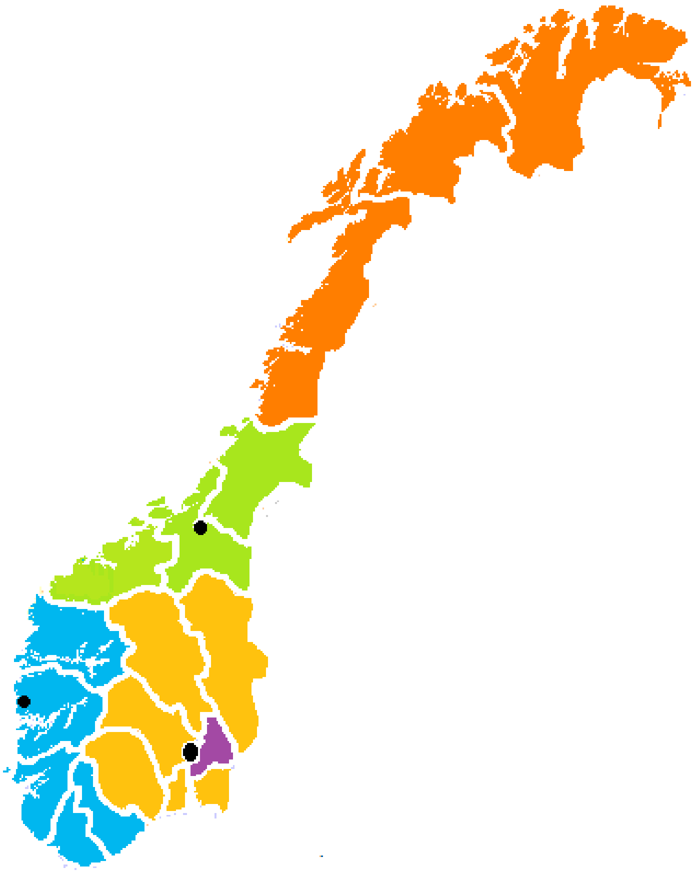

Table A2: Number of products by COICOP divisions and type.

COICOP division	Non-durab	Semi-durab	Durables	Services	Total
1 Food and non-alcoholic bevs	255				255
2 Alcoholic beverages, tobacco, ...	12				12
3 Clothing and footwear		97		5	112
4 Housing, water, electricity, ...	15				15
5 Furnishings, household equip	29	46	43	2	12
6 Health	39				39
7 Transport	4	7	9	17	37
8 Communication			7		7
9 Recreation and culture	27	27	23	4	81
11 Restaurants and hotels				38	38
12 Miscellaneous goods, services	43	7	4	6	6
Total	424	184	86	72	766

B The pooled distribution of normalized prices

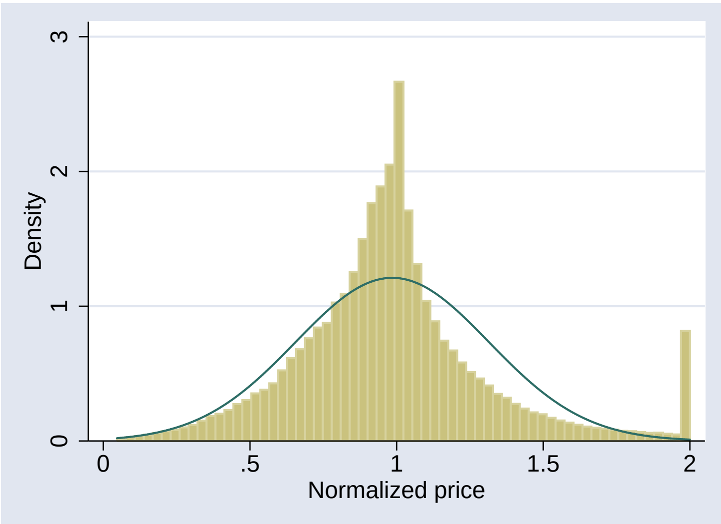

Figure B1: The pooled distribution of normalized prices truncated at 2.

C Price dispersion and the price level

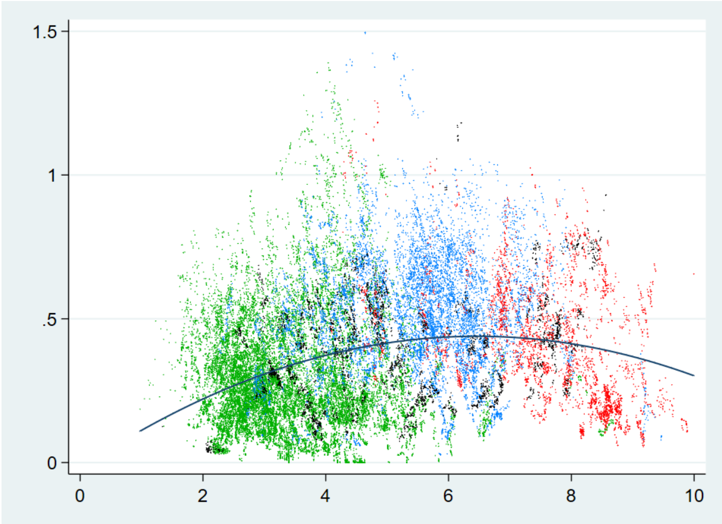

Figure C2: Scatter plot of the standard deviation (verical axis) and log of the mean price (horisontal axis) with a quadratic fitted line. Non-durables are green, semi-durables are blue, durables are red, and services are black.

D Regional variation

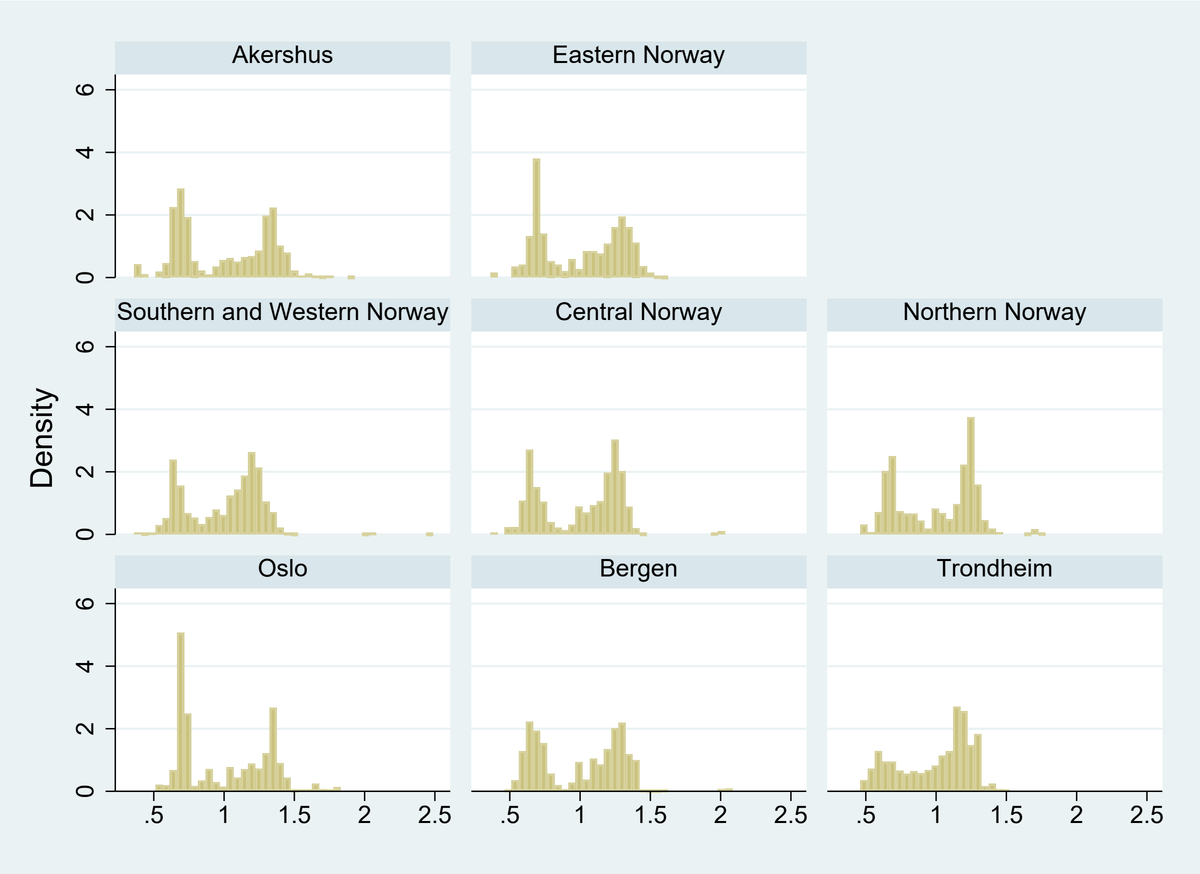

Figure D1: Regional distributions for the relative price of Coca Cola.

Table D1: Median dispersion in relative prices across regions.

		Standard deviation	
Region	N	Median	$(Q1 - Q3)$ range
Oslo	762	0.287	(0.169 – 0.471)
Bergen	765	0.300	(0.171 – 0.457)
Trondheim	762	0.319	(0.172 – 0.516)
Akershus county	766	0.308	(0.170 – 0.467)
Eastern Norway	766	0.306	(0.184 – 0.453)
Southern and Western Norway	766	0.331	(0.202 – 0.470)
Central Norway	764	0.291	(0.184 – 0.455)
Northern Norway	765	0.278	(0.167 – 0.416)

Note: N is number of products in each region.

E Robustness

Table E1: Descriptive statistics for relative price dispersion across products and over time. Balanced panel.

Dispersion measure	Median	$(Q1 - Q3)$ range
Standard deviation	0.319	(0.186 – 0.491)
IQR	0.308	(0.160 – 0.544)
$(P95 - P5)$ range	0.916	(0.524 – 1.447)

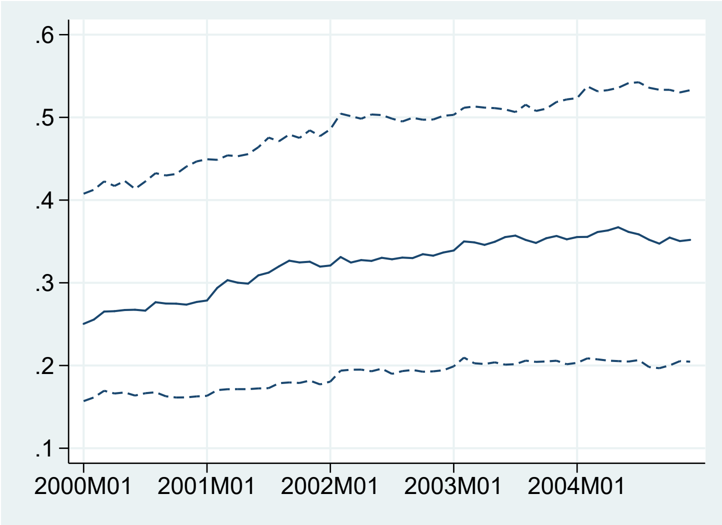

Figure E1: The first, second (median) and third quartiles of the standard deviations over time. Balanced panel.

F Store effects by COICOP division

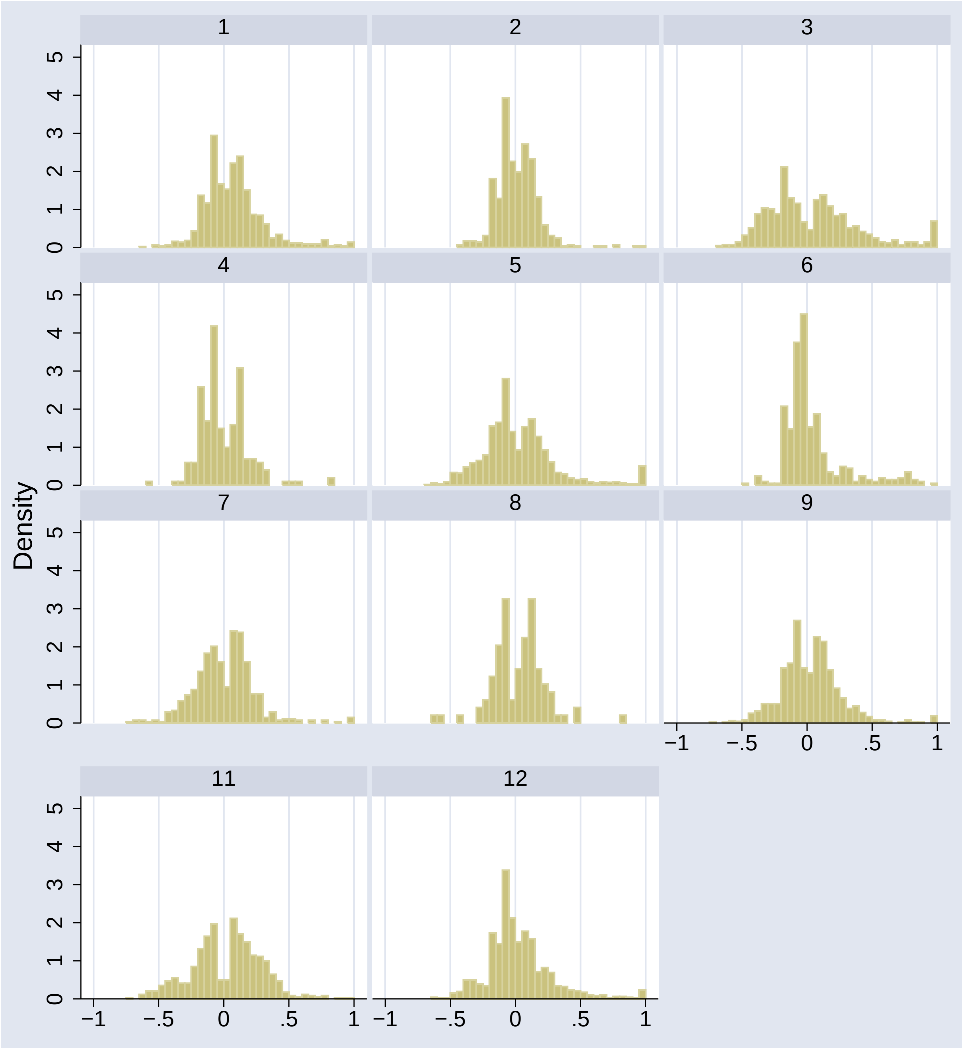

Figure F2: Histograms of the significant store effects (1 percent level significance) by COICOP *division. The histograms are truncated at 1.*

G Unconditional One Step Transition Probability Matrices

The elements in the matrix denote the probability of going from an initial quartile bin in period t (rows), to a destination quartile bin in period one, six and 12 months ahead.

Table G1: One Step Transition Probability Matrix, 1 month ahead. Means.

	Destination bin			
Initial bin	$QB1_{t+1}$	$QB2_{t+1}$	$QB3_{t+1}$	$QB4_{t+1}$
$QB1_t$	0.862	0.093	0.025	0.020
$QB2_t$	0.084	0.796	0.092	0.029
$QB3_t$	0.018	0.091	0.793	0.096
$QB4_t$	0.007	0.014	0.070	0.911

Table G2: One Step Transition Probability Matrix, 12 months ahead. Means.

	Destination bin			
Initial bin	$QB1_{t+12}$	$QB2_{t+12}$	$QB3_{t+12}$	$QB4_{t+12}$
$QB1_t$	0.604	0.233	0.093	0.071
$QB2_t$	0.181	0.496	0.224	0.091
$QB3_t$	0.060	0.203	0.498	0.237
$QB4_t$	0.031	0.063	0.184	0.726

Table G3: One Step Transition Probability Matrix, 6 months ahead. Median, mean and standard deviation.

	Destination bin			
Initial bin	$QB1_{t+6}$	$QB2_{t+6}$	$QB3_{t+6}$	$QB4_{t+6}$
$QB1_t$	0.717	0.176	0.046	0.020
	0.685	0.195	0.070	0.051
	(0.167)	(0.109)	(0.072)	(0.102)
$QB2_t$	0.152	0.609	0.162	0.037
	0.162	0.580	0.184	0.070
	(0.091)	(0.182)	(0.112)	(0.110)
$QB3_t$	0.031	0.173	0.614	0.157
	0.047	0.177	0.580	0.193
	(0.051)	(0.081)	(0.180)	(0.145)
$QB4_t$	0.012	0.030	0.146	0.801
	0.021	0.041	0.153	0.788
	(0.025)	(0.039)	(0.068)	(0.104)

H Conditional One Step Transition Probability Matrices

The conditional TPM is based on firms that have a nominal price change.

Table H1: One Step Conditional Transition Probability Matrix, 1 month ahead. Means.

	Destination bin			
Initial bin	$QB1_{t+1}$	$QB2_{t+1}$	$QB3_{t+1}$	$QB4_{t+1}$
$QB1_t$	0.692	0.174	0.076	0.057
$QB2_t$	0.150	0.578	0.191	0.079
$QB3_t$	0.057	0.162	0.582	0.194
$QB4_t$	0.028	0.050	0.144	0.780

Table H2: One Step Conditional Transition Probability Matrix, 12 month ahead. Means.

	Destination bin			
Initial bin	$QB1_{t+12}$	$QB2_{t+12}$	$QB3_{t+12}$	$QB4_{t+12}$
$QB1_t$	0.564	0.233	0.114	0.081
$QB2_t$	0.217	0.426	0.237	0.104
$QB3_t$	0.093	0.221	0.443	0.231
$QB4_t$	0.052	0.093	0.211	0.647

Table H3: One Step Conditional Transition Probability Matrix, 6 month ahead. Median, mean and standard deviation.

	Destination bin			
Initial bin	$QB1_{t+6}$	$QB2_{t+6}$	$QB3_{t+6}$	$QB4_{t+6}$
$QB1_t$	0.648	0.200	0.071	0.030
	0.627	0.207	0.096	0.065
	(0.174)	(0.104)	(0.090)	(0.108)
$QB2_t$	0.195	0.492	0.199	0.053
	0.207	0.481	0.212	0.089
	(0.113)	(0.176)	(0.118)	(0.109)
$QB3_t$	0.061	0.209	0.503	0.187
	0.080	0.208	0.495	0.210
	(0.077)	(0.096)	(0.167)	(0.137)
$QB4_t$	0.025	0.059	0.188	0.703
	0.038	0.074	0.190	0.698
	(0.042)	(0.062)	(0.086)	(0.130)

I Duration

Table I1: Fraction of stores within each quartile bin with monthly durations (mean, median and standard deviation).

Dur.	$QB1$	$QB2$	$QB3$	$QB4$	Dur.	$QB1$	$QB2$	$QB3$	$QB4$
1 mo.	0.208	0.244	0.244	0.175	7 mo.	0.030	0.034	0.033	0.029
	0.193	0.218	0.214	0.160		0.022	0.027	0.028	0.024
	0.111	0.124	0.131	0.095		0.033	0.035	0.031	0.027
2 mo.	0.116	0.151	0.150	0.102	8 mo.	0.031	0.028	0.029	0.028
	0.105	0.142	0.139	0.093		0.024	0.022	0.022	0.023
	0.074	0.085	0.085	0.062		0.033	0.030	0.032	0.030
3 mo.	0.095	0.110	0.108	0.081	9 mo.	0.025	0.025	0.026	0.026
	0.083	0.098	0.097	0.073		0.016	0.018	0.019	0.022
	0.066	0.069	0.067	0.053		0.034	0.029	0.029	0.025
4 mo.	0.077	0.085	0.087	0.070	10 mo.	0.037	0.030	0.033	0.047
	0.067	0.077	0.082	0.063		0.026	0.023	0.026	0.040
	0.059	0.051	0.052	0.046		0.041	0.033	0.035	0.039
5 mo.	0.088	0.083	0.082	0.083	11 mo.	0.022	0.025	0.024	0.024
	0.070	0.072	0.070	0.064		0.016	0.017	0.018	0.018
	0.086	0.071	0.070	0.076		0.029	0.035	0.027	0.024
6 mo.	0.050	0.052	0.054	0.052	12 mo.	0.219	0.134	0.134	0.283
	0.042	0.044	0.045	0.044		0.207	0.112	0.119	0.277
	0.043	0.041	0.036	0.043		0.144	0.118	0.113	0.149

Note: The table consists of the mean, the median and the standard deviations for each duration within a particular quartile.