Markups Across Space and Time*

Eric Anderson[†], Sergio Rebelo[‡], and Arlene Wong[§]

July 12, 2023

Abstract

We study the behavior of markups in the retail sector. Markups are measured with gross margins computed at the product level using the replacement cost for every item. We find that: (1) markups are relatively stable over time and mildly procyclical; (2) there is a large regional dispersion in markups; (3) regions with higher incomes have higher markups; (4) these higher markups do not result from less intense competition or regional differences in marginal costs; and (5) regional differences in markups are due to variations in assortment, not to deviations from uniform pricing practices. We propose a simple model consistent with these facts.

J.E.L. Classification: E30
Keywords: Gross margins, prices, marginal costs, business cycles.

*We thank Robert Barro, Gideon Borstein, Mark Bils, Joao Guerreiro, Laura Murphy, and Francisco Queirós for their comments and Miguel Santana, Martim Leitão, and Brian Livingston, and for excellent research assistance.

[†]Northwestern University.
[‡]Northwestern University, NBER, and CEPR.
[§]NBER.

1 Introduction

In this paper, we study the behavior of markups in the retail sector across different regions and time periods. We also propose a simple model consistent with the facts we document.

Our analysis of the regional behavior of markups uncovers four key facts. First, there is sizeable regional dispersion in markups. Second, regions with higher incomes and more expensive houses have higher markups. Third, these higher markups do not result from less competition or regional differences in marginal costs. Fourth, regional variation in markups occurs because retailers sell different goods in different regions. Consistent with the evidence in Della Vigna and Gentzkow (2019), we find that when the same item is available in different regions, it has a uniform price.

Our results regarding the temporal behavior of markups contribute to a long-standing debate in macroeconomics. We find that markups are remarkably stable over time and display a mild procyclical pattern. We also find that the conditional response of markups to monetary policy shocks and oil shocks is statistically insignificant.

Markups are notoriously difficult to measure because marginal costs are generally unobservable. Most empirical studies use structural approaches that rely on assumptions about production functions and market structure to infer marginal costs.[1] This literature, reviewed in depth by Nekarda and Ramey (2020), is divided in its conclusions about the cyclical properties of markups, in part because different studies rely on different structural assumptions.

We focus on the retail sector because its predominant variable cost, the cost of goods sold, can be used as a proxy for marginal cost. This cost accounts for over 80 percent of the total costs of retail firms. We use the gross margin (sales minus cost of goods sold as a fraction of sales) as a proxy for the markup.

Most of our analysis relies on scanner data from two major retailers—one based in the United States and the other in Canada. These datasets offer three significant advantages. First and foremost, they include the replacement cost for every item. This proxy for marginal

[1]For example, Bils (1987), Hall (1988), Galeotti and Schiantarelli (1998), Rotemberg and Woodford (1999), Bils and Kahn (2000), Kryvtsov and Midrigan (2012), and Bils, Klenow, and Malin (2018) infer the cyclicality of markups from the cyclicality of the cost shares of labor and other inputs. Hall (2014) bases his analysis on the cyclicality of advertising expenses. Collard-Wexler and De Loecker (2015) study the cyclical properties of markups implied by an estimated production function. Aguirregabiria (1999) explores the interaction between price and inventory decisions.

cost is the one used by the managers to make pricing decisions. Second, the data includes the price for every transaction rather than the average price across transactions. Third, the data covers stores in different regions, allowing us to compute regional markups.

For robustness, we also study the behavior of gross margins in the retail sector as a whole and in individual retail firms included in Compustat.

This paper is organized as follows. Section 2 briefly discusses the literature related to our paper. Section 3 describes the data we use. Section 4 contains our empirical findings. Section 5 discusses a simple model consistent with the facts we document. Section 6 concludes.

2 Literature review

In this section, we briefly discuss the literature on regional and temporal variations in markups that relates to our paper.

Regional variations in markups Our approach to estimating local business cycle effects is similar to that used by Coibion, Gorodnichenko, and Hong (2015), Beraja, Hurst, and Ospina (2019), and Stroebel and Vavra (2019). These authors study the response of prices to local business cycle conditions to infer the effect of monetary policy on aggregate fluctuations.

Our evidence is inconsistent with the models proposed by Greenhut and Greenhut (1975) and Thisse and Vives (1988) which predict that regions with higher markups have less intense competition.

The trade models based on non-homothetic preferences proposed by Bertoletti and Etro (2017) and Fajgelbaum, Grossman, and Helpman (2011) are consistent with the positive regional correlation between markups and income that we document. But both models make predictions that are inconsistent with our other findings. The Bertoletti and Etro (2017) model implies that variations in markups across regions are driven by deviations from uniform pricing. The Fajgelbaum, Grossman, and Helpman (2011) model implies that markups are countercyclical when costs are procyclical.

Temporal variation in markups Our results on the temporal behavior of markups contribute to answering a classic question in macroeconomics: are markups procyclical, acyclical, or countercyclical? The answer to this question is central to many key issues in macroeconomics, ranging from the slope of the Phillips curve and the size of the fiscal multiplier (Hall (2009)) to the cyclical movements of the share of labor in income (Kaplan and Zoch (2020)).

The presumption that markups are countercyclical has a long-standing tradition in macroeconomics. Notable examples of models featuring countercyclical markups include Rotemberg and Woodford's (1992) imperfect competition model, Ravn, Schmitt-Grohe, Uribe (2008)'s deep-habit model, Jaimovich and Floetotto (2008)'s entry and exit model, and models with sticky prices at the retail level and procyclical marginal cost. Examples of the latter class of environments include Golosov and Lucas (2007), Midrigan (2011), and the textbook New-Keynesian model (Woodford (2003) and Gali (2015)). As discussed in the introduction, we find no evidence in favor of the countercyclical markups featured in these models.

Instead, our time-series evidence favors models with sticky prices at the retail level and acyclical marginal costs (e.g., in Nakamura and Steinsson (2010), Coibion, Gorodnichenko, and Hong (2015) and Pasten, Schoenle, and Weber (2017)) and models with prices and wage rigidities at the manufacturing level (e.g., Erceg, Henderson, and Levin (2000), Christiano, Eichenbaum, and Evans (2005) and Christiano, Eichenbaum, and Trabandt (2016)). These models imply acyclical markups that are consistent with our evidence. Search models in which people devote time to search for lower prices generate procyclical markups because workers search less in expansions when the opportunity cost of search, the wage rate, is high (see, e.g., Alessandria (2009)). When this procyclicality is mild, these models are consistent with our time-series evidence.

Our regional evidence is inconsistent with most existing macroeconomic models because those models rely on homothetic preferences which result in markups that are independent of regional income.

Our paper significantly expands the existing body of evidence concerning retail markups. Prior work generally focuses on a single product category. For example, Dutta, Bergen, and Levy (2002) study orange juice, Nakamura and Zerom (2010) coffee, and Goldberg and Hellerstein (2012) beer.

Our research adds to the body of research that studies the empirical properties of retail prices. Notable examples of this literature include Bils and Klenow (2004), Nakamura and

Steinsson (2008), and Dhyne et al. (2006). Estimates from this literature are widely used to inform the calibration of macroeconomic models. Prominent examples of this approach include the models of monetary transmission proposed by Golosov and Lucas (2007) and Midrigan (2011), Mackowiak and Wiederholt's (2009) rational inattention model, as well as Beraja, Hurst, and Ospina's (2019) regional business cycles model.

3 Data

Our analysis focuses on the retail sector, which accounts for roughly 10 percent of aggregate employment. We use gross margins as proxies for markups. This approach is suitable for the retail industry because the cost of goods sold is the predominant variable cost. Gross margins might not be a good proxy for markups in other industries, such as manufacturing, where labor and other costs represent a larger fraction of total variable costs. Our primary data sources are two scanner data sets for a U.S. and a Canadian retail firm. To check robustness, we also study Compustat data for retail firms.

Firm level data The first data set, obtained from Compustat, includes quarterly panel data on sales, costs of goods sold, selling, general and administrative expenses, and net profits for retail firms from 1979 to 2014.[2] Our sample has 1,735 retail firms. The correlation between sales growth rates from Compustat data for the retail sector and sales growth rates from the U.S. Census Retail survey data is 70 percent.

Using Compustat data, we construct two margins for each firm f in quarter t:

$$(\text{Gross margin})_{ft} = \frac{\text{Sales}_{ft} - (\text{Cost of goods sold})_{ft}}{\text{Sales}_{ft}}, \tag{1}$$

$$\begin{aligned}(\text{Net operating profit margin})_{ft} &= \frac{\text{Sales}_{ft} - (\text{Cost of goods sold})_{ft} - (\text{Other expenses})_{ft}}{\text{Sales}_{ft}}, \\ &= (\text{Gross margin})_{ft} - \frac{(\text{Other expenses})_{ft}}{\text{Sales}_{ft}}.\end{aligned} \tag{2}$$

[2]The cost of goods sold does not include selling, general and administrative expenses. These expenses are reported separately from the cost of goods sold.

Other expenses include overhead expenses, rent, labor costs, and capital and property depreciation. For retail firms, these expenses are predominately fixed or quasi-fixed costs.

Large U.S. retailer data Our second data source is a scanner data set from a large retailer that operates over 100 stores in different U.S. states. This retailer sells grocery, health and beauty, and general merchandise products. We have weekly observations on quantities sold and retail and wholesale prices for each item in each store. An item is a good, defined by its stock-keeping unit code (SKU) in a particular store. We have roughly 3.6 million SKU-store pairs across 79 product categories. Our sample period begins in the 1st quarter of 2006 and ends in the 3rd quarter of 2009, so it includes the recession that started in the 4th quarter of 2007 and ended in the 2nd quarter of 2009.

Large Canadian retailer data Our third data source is a scanner data set from a large retailer that operates hundreds of stores in different Canadian provinces. This retailer sells products in 41 product groups, including clothing and footwear, toys, books, videos, and sporting and camping equipment. We have weekly observations on quantities sold and retail and wholesale prices for 15.6 million item-store pairs. The sample begins in the 1st quarter of 2016 and ends in the 4th quarter of 2018. The Canadian economy grew at a moderate pace during this period.

Our scanner data sets have two key features that distinguish them from several other scanner data sets.[3] First, they contain the price of every transaction instead of the average price across transactions. Second, the cost data measures the replacement cost, which is a good proxy for marginal cost. Moreover, the replacement cost is available at the store level rather than as a national average. This availability allows us to compute the gross margin as the difference between the price and the replacement cost for each item and store at each point in time.

Using these two scanner data sets, we construct the percentage gross margin for each item, i, at store s, in county k, at time t:

$$(\text{Gross margin})_{iskt} = \frac{\text{Price}_{iskt} - (\text{Replacement cost})_{iskt}}{\text{Price}_{iskt}}. \tag{3}$$

[3] Data from this retailer have been used in other studies, including Anderson, Jaimovich, and Simester (2015), McShane, Chen, and Anderson (2016), and Anderson, Malin, Nakamura, Simester, and Steinsson (2017).

Since the real GDP data we use to measure economic activity is available quarterly, we construct gross margins at a quarterly frequency by expenditure-weighting weekly gross margins.

We define the growth rate of the gross margin from $t-1$ to t for the subset of products that are in stock at time t and $t-1$ as:

$$g_{kt} \equiv \frac{\sum_s \sum_{i \in I_{i,t-1,t}} \omega_{iskt-1} \times \text{Gross margin}_{iskt}}{\sum_s \sum_{j \in I_{j,t-1,t}} \omega_{jskt-1} \times \text{Gross margin}_{jsk,t-1}},$$

where

$$\omega_{isk,t-1} = \frac{\text{Cost of goods sold}_{isk,t-1}}{\text{Total cost of goods sold}_{k,t-1}}.$$

and the cost of goods sold of an item is its replacement cost times the quantity sold.

We compute the chained gross margin as

$$\text{Gross margin}_{kt} = \prod_{d=1}^{t} g_{kd} \times \text{Gross margin}_{k0},$$

where Gross margin_{k0} denotes the weighted average of the gross margin in region k in period 0 computed using the cost of goods sold as weights. We use this measure of the gross margin, whose construction resembles the Laspeyres index, to study the margin cyclicality generated by changes in the margins of individual items. This measure abstracts from changes in margin resulting from product substitution between time $t-1$ and t.[4]

We also use data on the unemployment rate, real GDP growth, and estimates of monetary policy and oil price shocks. The monetary-policy shocks are identified from high-frequency Federal Funds futures data.[5] Oil-price shocks are identified using the approach proposed by Ramey and Vine (2010). The Appendix provides additional details on the process used to estimate these shocks.

4 Temporal variation in markups

This section documents the cyclical properties of gross margins, operating margins, sales, and cost of goods sold. We discuss the comovement and volatility of these series for the aggregate retail sector, at the firm and product level.

[4]We thank Mark Bils for suggesting that we use this measure of the gross margin.

[5]See Kuttner (2001) and Gurkaynak, Sack, and Swanson (2005) for details on the construction of these shocks.

4.1 Aggregate retail sector evidence

We construct aggregate measures of our variables for the retail sector using aggregate sales and aggregate costs. Table 1 summarizes the elasticity of different variables with respect to real GDP. This elasticity is estimated by regressing the year-on-year logarithmic difference of each variable on the year-on-year logarithmic difference of real GDP.

We see that gross margins are roughly acyclical or mildly procyclical. In contrast, sales and cost of goods sold are highly procyclical. These properties suggest that firms do not change markups in response to business-cycle fluctuations. Instead, the business cycle primarily affects quantities sold, operating profit margins, and the cost charged by suppliers, which is why sales and the cost of goods sold are highly procyclical.

Table 2 shows that gross margins are relatively stable compared to other variables. At a quarterly frequency, operating profit margins are 3.4 times more volatile than gross margins, while sales and costs are roughly 2.6 times more volatile than gross margins. The high volatility of operating profit margins compared to the volatility of gross margins suggests that fixed costs might be an important driver of profitability. Figure 1, which depicts the log differences from the prior year of gross margins and operating margins, illustrates the different volatility of these two variables.

4.2 Firm-level evidence

To study the cyclical properties of firm-level variables, we regress each variable on the year-on-year log-difference in real GDP using firm fixed effects. These fixed effects control for any permanent differences across firms, including differences in the degree of vertical integration between the retail and manufacturing operations.

Table 3 reports our elasticity estimates. The elasticity of the gross margin is small and statistically insignificant, while the elasticities of operating profits, sales, and cost of goods sold are positive and statistically significant. Consistent with the aggregate evidence, the firm-level evidence suggests that business cycles primarily affect costs and quantities sold rather than gross margins.

To study the volatility of a given variable at the firm level, we estimate the standard deviation of this variable for each firm and then compute the equally-weighted average of this statistic across firms. We report our results in Table 4. The operating profit margin is the most volatile variable in our sample, while the gross margin is the least volatile.

Nekarda and Ramey (2020) emphasize the importance of studying the conditional response of markups to various types of shocks. So, we use our firm-level data to study the conditional response of the gross margin and the operating profit margin to high-frequency monetary-policy shocks and oil-price shocks.[6] We estimate this response by running the following regression separately for the gross margin and the net operating profit margin:

$$\Delta \ln m_{it} = \beta_0 + \sum_k \beta_k \epsilon_{t-k} + \lambda_{q(t)} + \lambda_r + \eta_{it}, \tag{4}$$

where $\Delta \ln m_{it}$ is the year-on-year log-difference in the margin of firm i at time t. The variable ϵ_{t-k} is the aggregate shock at time $t-k$. The variables $\lambda_{q(t)}$, λ_r, and λ_i are fixed effects for the calendar quarter, recession, and firm.

Figure 2 depicts the implied impulse response functions. We see that the response of the gross margin is statistically insignificant for both monetary and oil-price shocks. In contrast, net operating profit margins fall in a statistically significant manner in response to both shocks.

4.3 Product-level evidence

There are two potential sources of measurement error in our aggregate data for the retail sector. First, gross margins are constructed using average costs instead of marginal costs. Second, changes in inventories can affect the cost of goods sold and potentially influence the cyclical properties of our empirical measure of the gross margin.[7] We now report results that are free of these two potential sources of measurement error. Our analysis is based on scanner data from two large retailers, one in the U.S. and the other in Canada. These data include transaction prices and replacement costs for every item. Using this information, we compute gross margins for every product in every store. We aggregate the weekly observations to construct monthly data.

[6] Our scanner data does not contain enough time-series observations to estimate the conditional response of the gross margin to shocks.

[7] Appendix A2 presents a version of our analysis where we adjust the cost of goods sold for changes in inventories. We still find that the elasticity of gross margins with respect to GDP is statistically insignificant.

4.4 Results for U.S. scanner data

In this section, we analyze product-level data for a major U.S. retailer. Figure 3 shows how the U.S. retailer reacted to the onset of the 2009 recession. This figure plots the distribution for gross margins, year-on-year log difference in sales and the number of unique items for the periods 2006-07 and 2008-09.

Each data point in the distribution is a region-quarter observation. For confidentiality reasons, we do not report the level of the average gross margin. In constructing Figure 3, we normalize the gross margins by subtracting the average gross margin for 2006-07 from the gross margins for 2006-07 and 2008-09. As a result, the normalized average gross margin for 2006-07 is zero.

We see that the regional distribution of the level of gross margins remained relatively stable with a slight shift to the left. In contrast, the distribution of year-on-year log difference in sales is more skewed in the Great Recession than in the 2006-07 period. This result is consistent with Bloom, Guvenen, and Salgado (2019), who find that sales growth becomes skewed during recessions.

The distribution of the number of unique items in each store shifted to the left during the recession. In other words, lower sales are associated with a smaller assortment and stable gross margins.

Table 5 reports the average, median, 10th, and 90th percentiles of the distribution of the three variables in Figure 3 for the expansion and recession periods. The gross-margin moments are similar across the two periods.[8] In contrast, the sales and number of items moments are all lower during the recession period.

To go beyond these unconditional moments, we compute the elasticity of the variables of interest with respect to the local unemployment rate and local real house prices.[9] Our approach is similar to that of Stroebel and Vavra (2019). We estimate the following regression:

$$\Delta \log \text{margins}_{mt} = \beta_0 + \beta_1 \Delta \log(Z_{mt}) + \varepsilon_{mt}, \tag{5}$$

[8]For confidentiality reasons, we do not report the average gross margin, only the difference in the average gross margin between the expansion and recession periods.

[9]We thank Emi Nakamura for sharing with us unemployment data for the regions in our scanner data.

where m denotes the region and t denotes the time period. We consider two possible alternative explanatory variables, Z_{mt}: the local unemployment rate and house prices instrumented with the housing supply elasticity proposed by Saiz (2010).[10] Since the Saiz (2010) instrument is static, for the regression with house prices, we consider the difference between the period 2005-2006 and 2007-08. For the regression with the unemployment rate, we consider the yearly log differences of the variables. The regression is estimated at the monthly frequency and includes region fixed effects.

Table 6 reports our results. The elasticity of the gross margin with respect to unemployment is statistically significant but very small (-0.003). The elasticity of the gross margin with respect to local house prices is statistically insignificant. The price and replacement cost elasticities with respect to unemployment are statistically significant but close to zero. Both metrics are statistically insignificant with respect to house prices. Sales elasticity is statistically significant and large for both the unemployment rate and local house prices, indicating that sales rise in periods when the local economy booms. Finally, the number of unique items carried in the store is negatively related to unemployment and positively related to house price, or procyclical.

Cost cyclicality One natural question is whether retail prices contribute to price inertia or simply reflect inertia in wholesale prices. To investigate this question, we divide products into three groups. The first group has acyclical costs, the second procyclical costs, and the third countercyclical costs. To classify costs according to their cyclicality, we regress the logarithmic change in the cost of goods sold on the difference in the local unemployment rate. We classify as procyclical (countercyclical) the cost of goods with a positive (negative) regression coefficient statistically significant at a 10 percent significance level. We classify as acyclical the cost of goods with an insignificant regression coefficient.

Table 7, Panel A shows that our findings about the acyclicality or mild procyclicality of gross margins hold regardless of whether costs are acyclical, procyclical, or countercyclical. Table 7, Panel B replicates this finding with a Canadian Retailer. We conclude that the

[10]This instrument uses information on the geography of a metropolitan area to measure the ease with which new housing can be built. The index assigns a high elasticity of housing supply to areas with a flat topology and without many water bodies, such as lakes and oceans. In low-elasticity areas, it is more difficult for the housing supply to respond to demand shocks, so these shocks produce larger movements in house prices.

behavior of retailers contributes to price inertia because, even for products with procyclical costs, retail margins are acyclical or mildly procyclical.

Passive and active margin changes To further investigate the cyclical properties of the gross margin, we divide margin changes into "passive" and "active." We define passive gross-margin changes for a given product as those that occur when the replacement cost of that product changes, but the company does not change the product's price. Active gross-margin changes for a given product result from price changes that occur regardless of whether the replacement cost has changed. We compute these changes at a daily frequency and then aggregate them at a monthly frequency using the average of the daily changes. Most (91 percent) margin changes are active.

Table 8 summarizes our results obtained using specification (5). We find that the probability of active margin change is acyclical.[11] This result holds both when we use the unemployment rate and local house prices as measures of local business conditions. Since changes in active margins are acyclical, overall margin changes are also acyclical.

Table 8 shows that, for both passive and active margin changes, the size of changes in gross margin and the changes in replacement cost conditional on changes in the gross margin are also acyclical. With respect to unemployment, the slope coefficients are statistically insignificant. With respect to house prices, the slope coefficients are statistically significant for passive margin chances but small in magnitude.

Figure 4 shows the histogram of the probability of passive margin changes, the size of gross margin changes, and the changes in replacement cost given margin changes during the expansion (2006-07) and recession periods (2008-09). Consistent with the notion that the probability of active margin changes is acyclical, we see that the two distributions are very similar. We also see that the dispersion of the size of active and passive margin changes is higher during the recession period.

Table 8 shows the standard deviation of year-on-year logarithmic changes in different variables. We see that gross margins, prices, and cost of goods sold are relatively stable. In contrast, sales and the number of unique items in the stores' assortment are volatile.

[11] Since the probability of passive margin changes is one minus the probability of active margin changes, the probability of passive margin changes is also acyclical.

4.5 Results for Canadian scanner data

We run regression (5) using the Canadian unemployment rate as an explanatory variable and region fixed effects, where a region is defined as a Census metropolitan area.[12] Table 10 reports our results. Recall that our data covers a period during which Canada experienced a moderate expansion. Quarterly real GDP growth rates ranged from 0.06 to 1.08 percent. While there is not much aggregate variation in growth rates, there is substantial regional variation. Our point estimates indicate that gross margins are slightly procyclical, but the gross-margin elasticity is statistically insignificant. We also find evidence that sales are strongly procyclical–the sales elasticity is negative (statistically significant at a 5 percent level) with respect to unemployment and positive with respect to oil prices. These results for a different country, set of goods, and cyclical period are broadly similar to those obtained for the U.S.

One advantage of the Canadian data is that changes in oil prices generate substantial regional variation in economic activity. Alberta, Saskatchewan, Newfoundland, and Labrador are all highly dependent on oil production. An unexpected rise in oil prices is a negative supply shock for all regions and a positive demand shock for oil-producing regions. In Table 10, we report estimates of β_2 obtained by running the following regression:

$$\Delta \log \text{margins}_{mt} = \beta_0 + \beta_1 \Delta \log(Z_t) + \beta_2 \Delta \log(Z_t) I_m + \theta I_m + \varepsilon_{mt}, \tag{6}$$

where m denotes the region (Census metropolitan area) and t denotes the time period. The variable Z_t denotes the oil price at time t. The variable I_m is equal to one if the region is a major oil producer and zero otherwise. The coefficient β_2 isolates the positive demand shock to oil-producing regions. We find that the gross margins are acyclical.

We now turn to the properties of active and passive margin changes. Active changes represent 93 percent of all gross margin changes. Table 10 reports our estimates of β_1 obtained using specification (5) and unemployment as a measure of cyclical conditions. This table also reports our estimates of β_2 obtained using specification (6) and changes in oil prices as the measure of cyclical conditions. As in our U.S. data set, we find that both the probability of active and passive changes in gross margins are acyclical.

[12] The Saiz (2010) instrument is not available for Canada, so we cannot run a version of regression (5) using house prices as an explanatory variable.

Table 12 shows the standard deviation of year-on-year logarithmic changes in different variables. As in our U.S. data set, gross margins, prices, and cost of goods sold are relatively stable. In contrast, sales and the number of unique items in stores' assortment are volatile.

4.6 Comparing with markups based on the Hall approach

In this subsection, we compare our markup estimates with alternative estimates obtained using the approach proposed by Hall (1986, 1988). Under the Hall approach, the markup estimate is a ratio where the numerator is the output elasticity with respect to a variable input, and the denominator is that input's cost share in total revenue.

In practice, data on output is often unavailable, so researchers proxy for output using sales revenues or value added, deflated with common industry-level price deflators. Bond et al. (2020) argue that this approximation can bias markup estimates. We use our data to show that there is indeed a sizeable bias from implementing the Hall approach with sales revenues instead of quantities.

We use our two scanner data sets to obtain firm-level markups using Hall's approach as follows. We first compute output by deflating sales using a price deflator calculated as the sales-weighted geometric average of product-level prices. We then regress output on measures of goods sold, labor, and capital. The quantity of goods sold is computed by deflating the cost of goods sold with a cost-weighted geometric average of product-level replacement costs. Labor is proxied by the number of employees.[13] Capital is estimated as book value deflated by the capital deflator for the retail industry.

Since input usage is correlated with the firm's unobserved productivity, we follow the approach suggested by Blundell and Bond (2000) and instrument goods sold with lagged goods sold. We run the regressions at a quarterly frequency. Since our capital measures are annual, we assume that capital is constant throughout the year.

Table 13 reports our results. The first column reports the elasticity of output with respect to the quantity of goods sold. This elasticity is close to one both for the U.S. and the Canadian firm. This finding is consistent with the quantity of goods sold being the predominant variable input in the retail industry. We can divide this elasticity by the share

[13] Our results are robust to using labor costs deflated by wages.

of the cost of goods sold in revenue to estimate the firm's markup. We do not report this estimate for confidentiality reasons. But we compare it, in lines 3 and 6, to the markup obtained using our approach. The ratio of the two markup estimates is 1.014 and 0.991 for the U.S. and Canadian firms, respectively.

To evaluate the quantitative impact of the bias emphasized by Bond et al. (2020), we compute the elasticity of *revenue* with respect to the quantity of goods sold. We then divide it by the share of the cost of goods sold in revenue to estimate the firm's markup. Column 3 of Table 13 shows that the elasticity of revenue with respect to the quantity of goods sold is much lower than the corresponding output elasticity. It is 0.848 versus 0.98 for the U.S. firm and 0.844 versus 0.873 for the Canadian firm. As a consequence, the implied markup is 14 percent (U.S. firm) and 13 percent (Canadian firm) lower than the one obtained using gross margins.

In sum, we find that the Hall approach implemented using output elasticities yields markup estimates that are very similar to those obtained using gross margins. This result increases our confidence in the reliability of the firm-level markups we estimate with Compustat data. We also find that the bias Bond et al. (2020) emphasize is large in our data. Implementing the Hall approach using revenue elasticities instead of output elasticities results in a roughly 14 percent decline in the estimated markup.

5 Regional variation in markups

Recall that in Figure 3 we showed that in the US retailer, the regional distribution is relatively similar in the Great Recession and in the expansion that preceded it. Each data point in the distribution is a region-quarter observation. The mean of the distribution is somewhat higher in the expansion period, which is consistent with the notion that margins are slightly procyclical. The same figure shows a large regional dispersion in the gross margins of our large retailer in both the expansion and the recession periods.

In this section, we use our scanner data to study the distribution of gross margins across regions. We can decompose the overall variance of the gross margins into a time series and a regional component. We denote by v_{mt} the gross margin of region m at time t, computed as a sales-weighted average of all items in stores located in this region. The variance of v_{mt}

is given by:

$$\begin{aligned}
var\left(v_{mt}\right) &= \frac{1}{TM-1}\sum_{t}\sum_{m}\left(v_{mt}-v\right)^{2} \\
&= \frac{1}{TM-1}\sum_{t}\sum_{m}\left(v_{mt}-v_{t}+v_{t}-v\right)^{2} \\
&\approx \frac{1}{T}\sum_{t}\underbrace{\frac{\sum_{m}\left(v_{mt}-v_{t}\right)^{2}}{M-1}}_{var_{t}(v_{m})}+\underbrace{\frac{\sum_{t}\sum_{m}\left(v_{t}-v\right)^{2}}{TM-1}}_{var(v_{t})}+2cov(v_{mt}-v_{t},v_{t}-v),
\end{aligned}$$

where T is the total number of time periods and M is the total number of regions. The variable v_t is the average gross margin across all regions at time t, computed as a sales-weighted average of all items in all stores. The variable v is the average of v_t across time. The variables $\frac{1}{T}\sum_t var_t\left(v_m\right)$ and $var\left(v_t\right)$ represent the average regional and time-series variance of gross margins, respectively. The variable $cov(v_{mt}-v_t, v_t-v)$ is the covariance between the time-series and the regional component.

Table 14 reports our results. The regional variance in gross margins, $\frac{1}{T}\sum_t var_t\left(v_m\right)$, is 0.103 while the time-series variation, $var(v_t)$, is 0.013. The covariance term, $cov(v_{mt}-v_t, v_t-v)$, is close to zero. This decomposition suggests that most of the variation in gross margins comes from the cross-section, not from the time series.

To study the source of regional variation in gross margins, we start with the following equation for the variance of gross margins across different markets conditional on period t, $var_t(v_m)$:

$$var_t(v_m) = var_t\left(\sum_j v_{jm}w_{jm}\right). \tag{7}$$

Here, v_{jm} is the gross margin of product j in market m and w_{jm} is the sales of product j in market m as a fraction of total sales in market m.

Expanding the terms on the right-hand side of equation (7), we obtain:

$$
\begin{aligned}
var_t(v_m) \quad = \quad & \underbrace{var_t\left[\sum_j (v_{jm} - \bar{v}_j)\bar{w}_j\right]}_{\text{differences in gross margins for the same item}} \\
& + \underbrace{var_t\left[\sum_j (w_{jm} - \bar{w}_j)\bar{v}_j\right]}_{\text{differences in assortment composition}} \\
& \underbrace{+var_t\left[\sum_j (v_{jm} - v_j)(w_{jm} - \bar{w}_j)\right]}_{\text{interaction term}} + \text{covariance terms.}
\end{aligned}
$$

The first term on the right-hand side of this equation measures the importance of differences in gross margins for the same item. This term is zero when there is uniform pricing, i.e., prices for the same product are identical across regions. The second term measures the importance of differences in assortment holding fixed the gross margin across regions. This term is zero when all regions have the same assortment composition. The third term measures the importance of the interaction between differences in assortment and differences in gross margins.

Table 15 reports the average over time of the components of this decomposition for the U.S. (panel A) and Canada (panel B). The first column of panels A and B reports results obtained using all items, including items sold in only some of the regions. In both panels, we find that the predominant driver of regional differences in gross margins is differences in assortment composition across regions. In contrast, regional differences in the gross margins of the same items account for very little of the regional variation in gross margins. In other words, when the same item is available in different regions, our retailers use roughly uniform pricing.

For robustness, we use our U.S. data to produce results obtained by restricting the sample to items sold in all regions.[14] We report these results in the second column of panel A. Here, the regional variation results from regional differences in consumer baskets. The results obtained using this restricted sample are similar to those obtained using the full sample.

[14]We do not report results for a sample of items sold in every market in Canada because the number of such items is relatively small.

We now investigate which variables might explain the regional variation in gross margins. Column 1 of Table 16 shows that gross margins in our U.S. data are positively correlated with measures of income or wealth. These measures include the logarithm of household income and the logarithm of the median house value. In contrast, gross margins are uncorrelated with a measure of competition (the Herfindahl index) and a proxy for higher transportation costs (a dummy variable that takes the value one for counties classified by the census as rural).

We find that there is indeed a positive cross-sectional correlation between local income and local gross margins. But these differences in gross margins across regions are explained by differences in assortment, not by deviations from uniform pricing. These results are consistent with the evidence in Della Vigna and Gentzkow (2019). They are also consistent with recent work by Neiman and Vavra (2018) that shows that households concentrate their spending on different goods. We add to these results by providing direct evidence of differences in gross margins and assortment across regions.

Column 2 of Table 16 shows that gross margins in our Canadian data are also positively correlated with measures of income or wealth. These measures include the logarithm of household income and the logarithm of the median house value. We also find a positive correlation between the unique number of items sold in a region and regional household income. This correlation is 0.42 for the Canadian retailer and 0.17 for the U.S. retailer.

6 A simple model

In this section, we present a simple model that is consistent with the key facts we document: (1) markups are relatively stable over time and mildly procyclical; (2) there is a large regional dispersion in markups; (3) regions with higher incomes have higher markups; (4) these higher markups do not result from less intense competition or regional differences in marginal costs; and (5) regional differences in markups are due to variations in assortment, not to deviations from uniform pricing practices.

We use our model to study how markups vary over time and across regions that differ in productivity. A key feature of the model that allows it to be consistent with our regional evidence is that household preferences are non-homothetic.

To simplify the notation, we omit region subscripts. Additionally, when it doesn't compromise clarity, we also omit time subscripts.

Households Each region has a continuum of measure one of identical households. Households supply exogenously N units of labor and decide how much to consume of a homogeneous good, z. In addition, they buy one unit of each variety $i \in (0,1)$ and choose the quality of each variety, q_i. The household's lifetime utility is given by,

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u_t,$$

where u_t denotes momentary utility. The symbol E_0 denotes the expectation conditional on information available at time zero.

As in Melitz and Ottaviano (2008), we assume that momentary utility has a non-homothetic, quadratic form,

$$u_t = z_t + \alpha \int_0^1 q_{i,t} di - \frac{\gamma}{2} \int_0^1 q_{i,t}^2 di - \frac{\eta}{2} \left(\int_0^1 q_{i,t} di \right)^2,$$

where α, γ and η are all positive. The parameters α and η control the patterns of substitution between the homogeneous good and the differentiated varieties. The parameter γ controls the degree of differentiation between varieties. When $\gamma = 0$, the different varieties are perfect substitutes.

To simplify, we assume that the differentiated goods are produced in the region where they are consumed and that the associated monopolist profits, $\pi_i, i \in (0,1)$, are distributed to the representative household. The household budget constraint is

$$z + \int_0^1 p_i q_i di = wN + \int_0^1 \pi_i di + \pi_z,$$

where w is the wage rate, π_z is the profit of the producers of the homogeneous good. The price of one unit of variety i is linear in quality, p_i is the price per unit of quality of good i of one unit of good i. We choose good z as the numeraire, so its price equals one.

The first-order conditions for this problem are

$$\lambda = 1,$$

$$\alpha - \gamma q_i - \eta \left(\int_0^1 q_i di \right) = \lambda p_i.$$

The implied demand function has a linear form,

$$p_i = \alpha - \gamma q_i - \eta \left(\int_0^1 q_i di \right).$$

The absolute value of the own elasticity of demand evaluated in a symmetric equilibrium is

$$\left| \frac{dq_i}{dp_i} \frac{p_i}{q_i} \right| = \frac{\alpha/q_i - \gamma - \eta}{\gamma} > 0.$$

Since, in equilibrium, the price is always positive, the numerator of this expression is positive When q_i increases, the absolute value of the demand elasticity falls. In other words, demand becomes less elastic. In equilibrium, this property leads to a higher markup.

Homogeneous good producers It takes one unit of labor to produce one unit of good z. The problem of the competitive, homogeneous good producers is to maximize profits given by,

$$\pi_z = z - wz.$$

The first-order condition for this problem is

$$w = 1.$$

At the optimum, $\pi_z = 0$.

Monopolist problem Producing one unit of good i with quality q_i costs q_i/A units of labor. The profit of monopolist i is

$$\pi_i = p_i q_i - w \frac{q_i}{A}.$$

The first-order condition for this problem is,

$$\alpha - 2\gamma q_i - \eta \left(\int_0^1 q_i di \right) = \frac{w}{A}.$$

Labor market The labor market clearing condition is

$$\frac{1}{A} \int_0^1 q_i di + z = N.$$

The first term on the right-hand side is the labor employed in producing the continuum of measure one differentiated goods. The second term is the labor used to produce the homogeneous good.

Equilibrium We impose the regularity condition $\alpha > 1/A$ so that the quality consumed of differentiated goods is positive. The following proposition summarizes the properties of the equilibrium.

Proposition 1. *The equilibrium quality, price, and markup of each differentiated good i are*

$$q_i = \frac{1}{2\gamma + \eta}\left(\alpha - \frac{1}{A}\right),$$

$$p_i = \alpha - \frac{\gamma + \eta}{2\gamma + \eta}\left(\alpha - \frac{1}{A}\right),$$

$$\frac{p_i}{1/A} = \alpha A \frac{\gamma}{2\gamma + \eta} + \frac{\gamma + \eta}{2\gamma + \eta}.$$

The equilibrium level of consumption of the homogeneous good is,

$$z = N - \frac{1}{2\gamma + \eta}\frac{1}{A}\left(\alpha - \frac{1}{A}\right).$$

When $\gamma = 0$, varieties are perfect substitutes and their price per quality unit equals marginal cost ($1/A$).

Consider an economy in which A and N vary over time. The markup is procyclical with respect to A. The elasticity of the markup with respect to A is less than one, so there is incomplete passthrough from cost to price. The markup is acyclical with respect to N. Suppose that the business cycle is driven by a mixture of shocks to A and N. In this case, the markup is mildly procyclical.

The homogeneous good z has a uniform price across all regions. Regional variation in the prices of differentiated goods occurs because higher productivity regions choose higher-quality goods, and these goods have higher markups. This implication is consistent with our finding, reported in Table 13, that the predominant driver of regional differences in gross margins is differences in assortment across regions.

An interesting property of this model is that it is consistent with the evidence provided by Jaimovich, Rebelo, and Wong (2019). In response to a decline in A (a negative cost shock) households trade down, that is, they buy goods of lesser quality and relatively lower price.

7 Conclusion

The Dixit-Stiglitz model of monopolistic competition, which features a constant demand elasticity, has become the workhorse of modern macroeconomics. The regional facts we document in this paper suggest that we might need to move away from this framework and study models where the demand elasticity varies with the level of consumption. Such a shift could have significant implications.

In this paper, we take an initial step in this direction. We show that the linear demand system employed by Melitz and Ottaviano (2008) in a trade context can generate demand patterns that are consistent with the behavior of retail markups across space and time.

8 References

Aguirregabiria, Victor. "The Dynamics of Markups and Inventories in Retailing Firms." *The Review of Economic Studies* 66, no. 2 (1999): 275-308.

Alessandria, G., "Consumer search, price dispersion, and international relative price fluctuations," *International Economic Review*, 50(3), pp.803-829, 2009.

Anderson, Eric, Benjamin A. Malin, Emi Nakamura, Duncan Simester, and Jón Steinsson "Informational Rigidities and the Stickiness of Temporary Sales," *Journal of Monetary Economics*, 90, 64-83, October 2017.

Anderson, Eric. Nir Jaimovich and Duncan Simester "Price Stickiness: Empirical Evidence of the Menu Cost Channel," *Review of Economics and Statistics*, vol. 97, issue 4, 813-826, October 2015.

Beraja, Martin, Erik Hurst and Juan Ospina "The Aggregate Implications of Regional Business Cycles," *Econometrica*, 87, no. 6: 1789-1833. 2019.

Bertoletti, Paolo and Federico Etro, "Monopolistic Competition when Income Matters," *The Economic Journal*, 127, 1468-0297, 2017.

Bils, Mark "The Cyclical Behavior of Marginal Cost and Price," *The American Economic Review*, 77(5): 838–55, 1987.

Bils, Mark, and James A. Kahn "What Inventory Behavior Tells Us About Business Cycles." *The American Economic Review*, 90(3): 458–81, 2000.

Bils, Mark, and Peter J. Klenow "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy*, 112(5): 947–985, 2004.

Bils, Mark, and Peter J. Klenow "Quantifying Quality Growth, " *The American Economic Review*, 91(4): 1006-1030, September 2011.

Bils, Mark, Peter J. Klenow, and Benjamin A. Malin "Resurrecting the Role of the Product Market Wedge in Recessions," *The American Economic Review* 108(4-5), 1118-46, 2018.

Bloom, Nicholas, Fatih Guvenen and Sergio Salgado "Skewed Business Cycles," National Bureau of Economic Research working paper No. w26565, 2019

Bond, Steve Arshia Hashemi, Greg Kaplan and Piotr Zoch "Some Unpleasant Markup Arithmetic: Production Function Elasticities and their Estimation from Production Data, " *Journal of Monetary Economics*, 121, 1-14, 2021.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans "Nominal Rigidities and the Dynamic Effects of A Shock to Monetary Policy," *Journal of Political Economy*, 113(1): 1–45, 2005.

Christiano, Lawrence J., Martin Eichenbaum, and Mathias Trabandt, "Unemployment and Business Cycles," *Econometrica*, 84 (4), 1523–1569, 2016.

Cochrane, John, and Monika Piazzesi, "The Fed and Interest Rates: A High-Frequency Identification," *The American Economic Review*, 92(2), pp.90-95, 2002.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong, "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications," *The American Economic Review* 105 (3): 993-1029, 2015.

Collard-Wexler, Allan and Jan De Loecker "Reallocation and Technology: Evidence from the US Steel Industry," *The American Economic Review,* Vol. 105, No. 1, pp. 131-171, January 2015.

Della Vigna, Stefano and Matthew Gentzkow "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics* 134, no. 4: 2011-2084, 2019.

Dhyne, Emmanuel, Luis J. Alvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lunnemann, Fabio Rumler, and Jouko Vilmunen. "Price changes in the euro area and the United States: Some facts from individual consumer price data," *Journal of Economic Perspectives* 20, no. 2 (2006): 171-192.

Dutta, Shantanu, Mark Bergen, and Daniel Levy. "Price flexibility in channels of distribution: Evidence from scanner data," *Journal of Economic Dynamics and Control* 26, no. 11 (2002): 1845-1900.

Eichenbaum, Martin, Nir Jaimovich and Sergio Rebelo, "Reference Prices, Costs, and Nominal Rigidities," *The American Economic Review*, Vol. 101, No. 1, pp. 234-262, February 2011.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics*, vol. 46, no. 2, 281-314, 2000.

Faber, Benjamin and Thibault Fally "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," forthcoming *Review of Economic Studies*, 2020.

Fajgelbaum, Pablo, Gene M. Grossman and Elhanan Helpman "Income Distribution, Product Quality, and International Trade," *Journal of Political Economy*, Vol. 119, No. 4, pp. 721-765, August 2011.

Galeotti, M., and E Schiantarelli, "The Cyclicality of Markups in a Model with Adjustment Costs: Econometric Evidence for U.S. Industry," *Oxford Bulletin of Economics and Statistics* 60:121-142, 1998.

Gali, Jordi *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press, 2015.

Goldberg, Pinelopi, and Rebecca Hellerstein "A structural approach to identifying the sources of local currency price stability," *The Review of Economic Studies* 80, no. 1 (2013): 175-210.

Golosov, Mikhail and Robert E. Lucas Jr., "Menu Costs and Phillips Curves," *Journal of Political Economy*, vol. 115, pages 171-199, 2007.

Gorodnichenko, Yuriy, and Michael Weber, “Are Sticky Prices Costly? Evidence from the Stock Market,” *The American Economic Review* 106(1): 165-99, 2016.

Greenhut, John G. and Melvin L. Greenhut “Spatial Price Discrimination, Competition and Locational Effects. ” *Economica*, 42(168), pp.401-419, 1975.

Gürkaynak, Refet S., Brian P. Sack, and Eric T. Swanson “Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements,” *International Journal of Central Banking*, 1(1): 55–93, 2005.

Hall, Robert E. “The Relation between Price and Marginal Cost in U.S. Industry.” *Journal of Political Economy* 96 (October): 921–47, 1988.

Hall, Robert E. “What the Cyclical Response of Advertising Reveals about Markups and other Macroeconomic Wedges.” manuscript, Stanford University, 2014.

Hall, Robert E. “By How Much Does GDP Rise If the Government Buys More Output?,” *Brookings Papers on Economic Activity*, The Brookings Institution 40: 2, 2009.

Jaimovich, Nir and Max Floetotto “Firm Dynamics, Markup Variations, and the Business Cycle,” *Journal of Monetary Economics*, Volume 55, Issue 7, Pages 1238-1252, October 2008.

Jaimovich, Nir Sergio Rebelo and Arlene Wong “Trading Down and the Business Cycle,” *Journal of Monetary Economics* 102: 96-121, 2019.

Jaimovich, Nir Sergio Rebelo, Arlene Wong, and Miao Ben Zhang “Trading Up and the Skill Premium,” *NBER Macroeconomics Annual*, 34(1), pp. 285-316, 2019.

Kaplan, Greg and Piotr Zoch “Markups, Labor Market Inequality and the Nature of Work,” manuscript, University of Chicago, 2020.

Karabarbounis, Loukas “The Labor Wedge: MRS vs. MPN,” *Review of Economic Dynamics*, 17(2), 206-223, 2014.

Karabarbounis, Loukas and Brent Neiman “Accounting for Factorless Income,” manuscript, Federal Reserve Bank of Minneapolis, 2018.

King, Robert G., Charles I. Plosser, and Sergio T. Rebelo. ”Production, growth and business cycles: II. New directions.” Journal of Monetary Economics 21, no. 2-3 (1988): 309-341.

Kryvtsov, Oleksiy, and Virgiliu Midrigan. "Inventories, markups, and real rigidities in menu cost models." *Review of Economic Studies* 80, no. 1: 249-276, 2013.

Kuttner, Kenneth "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market," *Journal of Monetary Economics*, 47: 523–544, 2001.

Mackowiak, Bartosz, and Mirko Wiederholt. "Optimal sticky prices under rational inattention," *American Economic Review*, 99, no. 3 (2009): 769-803.

McShane, Blakeley B., Chaoqun Chen, Eric T. Anderson, and Duncan I. Simester "Decision Stages and Asymmetries in Regular Retail Price Pass-through." *Marketing Science*, 35(4), pp. 619-639, 2016.

Melitz, Marc, and Gianmarco Ottaviano. "Market size, trade, and productivity, " *The Review of Economic Studies* 75, no. 1 (2008): 295-316.

Midrigan, Virgiliu "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *Econometrica*, 79(4), pages 1139-1180, July 2011.

Nakamura, Emi, and Jón Steinsson. "Five facts about prices: A Reevaluation of Menu Cost Models." *The Quarterly Journal of Economics* 123, no. 4 (2008): 1415-1464.

Nakamura, Emi and Jón Steinsson "Monetary Non-Neutrality in a Multisector Menu Cost Model," *Quarterly Journal of Economics*, pages 961-1013, August 2010.

Nakamura, Emi and Jón Steinsson "High Frequency Identification of Monetary Non-Neutrality: The Information Effect," *Quarterly Journal of Economics*, 133, no. 3: 1283-1330, 2018.

Nakamura, Emi, and Dawit Zerom "Accounting for Incomplete Pass-through," *The Review of Economic Studies* 77, no. 3 (2010): 1192-1230.

Neiman, Brent and Joseph Vavra "The Rise in Household Spending Concentration," Working Paper, SSRN, March 2018.

Nekarda, Christopher and Valerie A. Ramey, "The Cyclical Behavior of the Price-Cost Markup," forthcoming, *Journal of Money, Credit, and Banking*, 2020.

Pasten, Ernesto, Raphael Schoenle, Michael Weber, "Price Rigidities and the Granular Origins of Aggregate Fluctuations," NBER Working Paper No. 23750, August 2017.

Ramey, Valerie A. and Daniel J. Vine "Oil, Automobiles, and the U.S. Economy: How Much Have Things Really Changed?," *NBER Macroeconomics Annual*, Vol. 25, No. 1, pp. 333-368, 2010.

Ravn, Morten Stephanie Schmitt-Grohé, and Martín Uribe "Deep Habits," *The Review of Economic Studies*, Volume 73, Issue 1, Pages 195–218, January 2006.

Rotemberg, Julio J., and Michael Woodford. "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity." *Journal of Political Economy* 100, no. 6 (1992): 1153-1207.

Rotemberg, Julio J., and Michael Woodford "The Cyclical Behavior of Prices and Costs." In *Handbook of Macroeconomics.* , ed. John B. Taylor and Michael Woodford, 1051–135. Elsevier, 1999.

Saiz, Albert "The Geographical Determinants of Housing Supply," *The Quarterly Journal of Economics*, Vol. 125, No. 3, pp. 1253-1296, August 2010.

Smets, Frank, and Rafael Wouters."Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American Economic Review* 97, no. 3 (2007): 586-606.

Stroebel, Johannes and Joseph Vavra "House Prices, Local Demand and Retail Prices," *Journal of Political Economy*,127, no. 3: 1391-1436, 2019.

Thisse, Jacques-Francois, and Xavier Vives "On the Strategic Choice of Spatial Price Policy. " *The American Economic Review*, pp.122-137, 1988.

Traina, James "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," manuscript, University of Chicago, 2018.

Woodford, Michael *Interest and Prices: Foundations of A Theory of Monetary Policy.* Princeton, N.J.:Princeton University Press, 2003.

A Appendix

A.1 Monetary policy and oil shocks

In section 3.2, we study the conditional response of firms' gross and net operating margins to high-frequency monetary policy shocks and oil-price shocks. This appendix discusses how these shocks are identified.

Monetary policy shocks are identified using high-frequency data on the Federal Funds futures contracts. This approach has been used by Kuttner (2001), Cochrane and Piazessi (2002), Nakamura and Steinsson (2018), Gorodnichenko and Weber (2016), and others. The future rate reflects the market expectations of the average effective Federal Funds rate during that month. It therefore provides a market-based measure of the anticipated path of the Federal Funds rate.

A current period monetary policy shock is defined as:

$$\epsilon_t = \frac{D}{D-t}\left(ff^0_{t+\triangle^+} - ff^0_{t-\triangle^-}\right) \tag{8}$$

where t is the time when the FOMC issues an announcement, $ff^0_{t+\triangle^+}$ is the Federal Funds futures rate shortly after t, $ff^0_{t-\triangle^-}$ is the Federal Funds futures rate just before t, and D is the number of days in the month. The $D/(D-t)$ term adjusts for the fact that the Federal Funds futures settle on the average effective overnight Federal Funds rate.

We consider a 60-minute time window around the announcement that starts $\triangle^- = 15$ minutes before the announcement. Examining a narrow window around the announcement ensures that the only relevant shock during that time period (if any) is the monetary policy shock. Following Cochrane and Piazessi (2002) and others, we aggregate up the identified shocks to obtain a quarterly measure of the monetary policy shock.

Oil-price shocks are identified using the approach proposed by Ramey and Vine (2010), updated to the recent period. We estimate a VAR system with monthly data

$$Y_t = A(L)Y_{t-1} + U_t.$$

The vector Y_t includes the following variables (in order): nominal price of oil, the CPI, nominal wages of private production workers, industrial production, civilian hours, and the federal funds rate. The function $A(L)$ is a matrix of polynomials in the lag operator L, and U is a vector of disturbances. All variables, except the federal funds rate, are in logs. We include a linear time trend and 6 lags of the variables. The shock to oil prices is identified using a standard Cholesky decomposition. The shocks are aggregated to a quarterly frequency to match the frequency of our firm level data.

A.2 Correcting gross margins for changes in inventories

One potential source of measurement error in our aggregate retail and firm level data stems from the possibility that the cost of goods sold might reflect goods purchased in previous periods and stored as inventory. As a result, the cost of goods sold does not measure the true marginal replacement cost.

We deal with this issue in Section 3.4 by using actual replacement cost for a retailer. Here, we use instead a perpetual inventory approach to correct the cost of goods sold for changes in inventories.

Denote by $\bar{C}_t$ the observed cost of goods sold and by C_t the true cost of goods sold. The observed cost of goods sold is

$$\bar{C}_t = \alpha_t \bar{C}_{t-1} + (1 - \alpha_t) C_t,$$

where

$$\alpha_t = \frac{\text{Starting period inventories}_t}{\text{Sales}_t}.$$

We assume that if $\alpha_t \geq 1$, then

$$\bar{C}_t = C_t/(1 + \pi_t),$$

where π_t is the rate of change in the producer price index for final goods from the Bureau of Labor Statistics. This equation implies that, if the inventories at the start of the period exceed sales in that period, then the goods sold in that period come from inventories.[15] The observed value of cost of good sold is then assumed to be given by the true cost of goods sold, deflated by the producer price index.

[15]This occurrence is rare, particularly at the annual frequency. The average retailer ratio of inventories to sales is about 12%.

The true cost of goods sold is given by

$$C_t = \frac{\bar{C}_t - \alpha_t \bar{C}_{t-1}}{1 - \alpha_t}, \quad \text{if } \alpha_t < 1$$

and

$$\bar{C}_t = C_t/(1 + \pi_t), \quad \text{if } \alpha_t \geq 1.$$

We assume as starting value $\bar{C}_0 = C_0$ and implement our approach separately for each firm.

The gross margin adjusted for changes in inventories is given by

$$\frac{\text{Sales}_t - \text{C}_t}{\text{Sales}_t}.$$

We use this adjusted measure to re-estimate the elasticity of gross margins with respect to real GDP. We regress the year-on-year logarithmic difference of each variable on the year-on-year logarithmic difference of real GDP.

Table 15 shows our results from Section 3, which do not adjust for inventories, as well as the elasticities estimated using gross margins adjusted for changes in inventories. We see that while point estimates are different, the elasticity of gross margins with respect to GDP growth remain statistically insignificant when we use the adjusted gross-margin measures.

Tables and Graphs

Table 1: Cyclicality of Aggregate Retail Trade Variables

	Elasticity wrt GDP			
	Quarterly		Annual	
	Coefficient	Standard error	Coefficient	Standard error
Gross margins	0.162	(0.256)	0.376	(0.616)
Operating profit margins	2.286**	(0.895)	5.233	(3.632)
Sales	8.089***	(0.45)	9.279***	(1.976)
Cost of goods sold	8.104***	(0.43)	9.140***	(2.154)

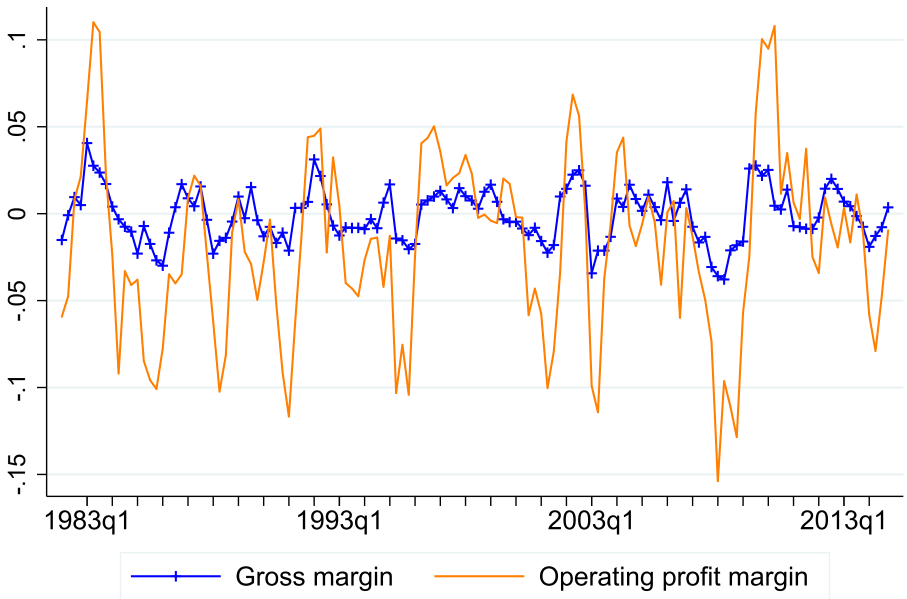
Notes: Variables are log-difference from the prior year. Data is from Compustat and the BLS. Each row is estimated from a separate regression of the dependent variables on GDP, as described in Section 4.1. We estimate the elasticities at quarterly and annual frequencies using data from 1980 to 2013. There are 136 and 33 observations for the quarterly and annual frequency regressions, respectively. Standard errors are in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels, respectively.

Table 2: Volatility of Aggregate Retail Trade Variables

Standard Deviation	Quarterly	Annual
Gross margins	0.017	0.011
Operating profit margins	0.057	0.051
Sales	0.046	0.062
Cost of goods sold	0.045	0.060

Notes: Variables are log-difference from the prior year. Data is from Compustat and the BLS. Standard deviations are computed at quarterly and annual frequencies using data from 1980-2013.

Figure 1: Time-series of Aggregate Retail Trade Variables



Notes: Variables are log-difference from the prior year. Data is from Compustat and the BLS. The data is plotted at a quarterly frequency.

Table 3: Cyclicality of Firm-Level Variables

	Elasticity wrt GDP			
	Quarterly		Annual	
	Coefficient	Standard error	Coefficient	Standard error
Gross margins	0.31	(0.37)	0.15	(0.55)
Operating profit margins	3.03***	(0.96)	3.60***	(1.11)
Sales	3.18***	(0.32)	3.64***	(0.67)
Cost of goods sold	3.09***	(0.32)	3.58***	(0.70)

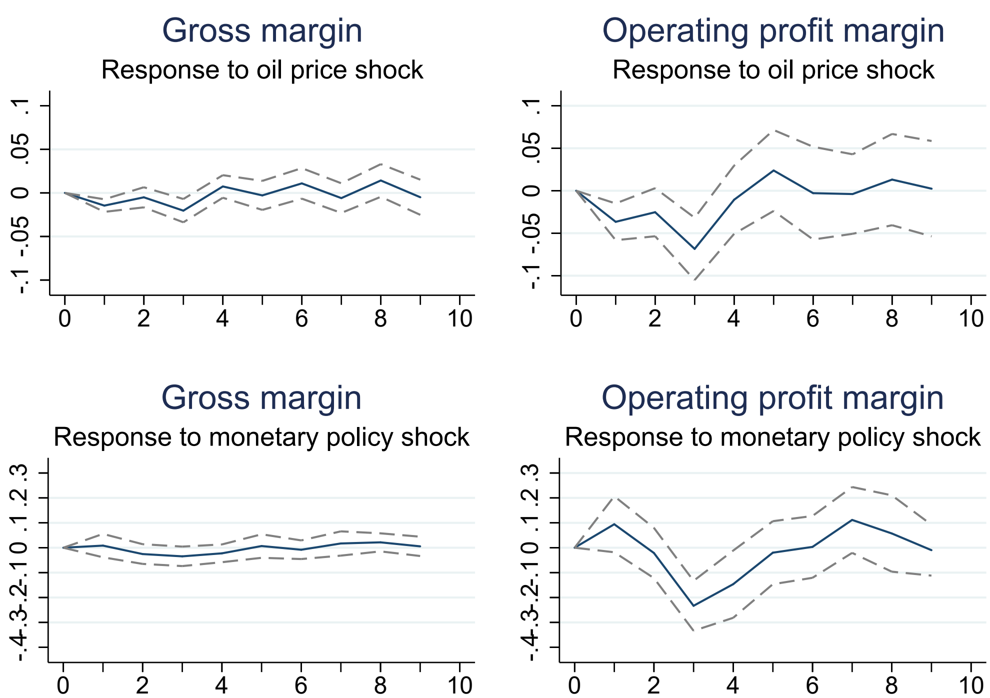
Notes: Variables are log-difference from the prior year. Data is from Compustat and the BLS. Each row is estimated from a separate regression of the dependent variables on GDP, including firm fixed effects, as described in Section 4.2. We estimate the elasticities at quarterly and annual frequencies using data from 1980-2013. 48,423 (10,312) observations are used in the estimation of the margins regression at quarterly (annual) frequency. Standard errors are in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table 4: Volatility of Firm-Level Variables

Standard Deviation	Quarterly	Annual
Gross margins	0.061	0.480
Operating profit margins	0.254	0.699
Sales	0.080	0.364
Cost of goods sold	0.084	0.407

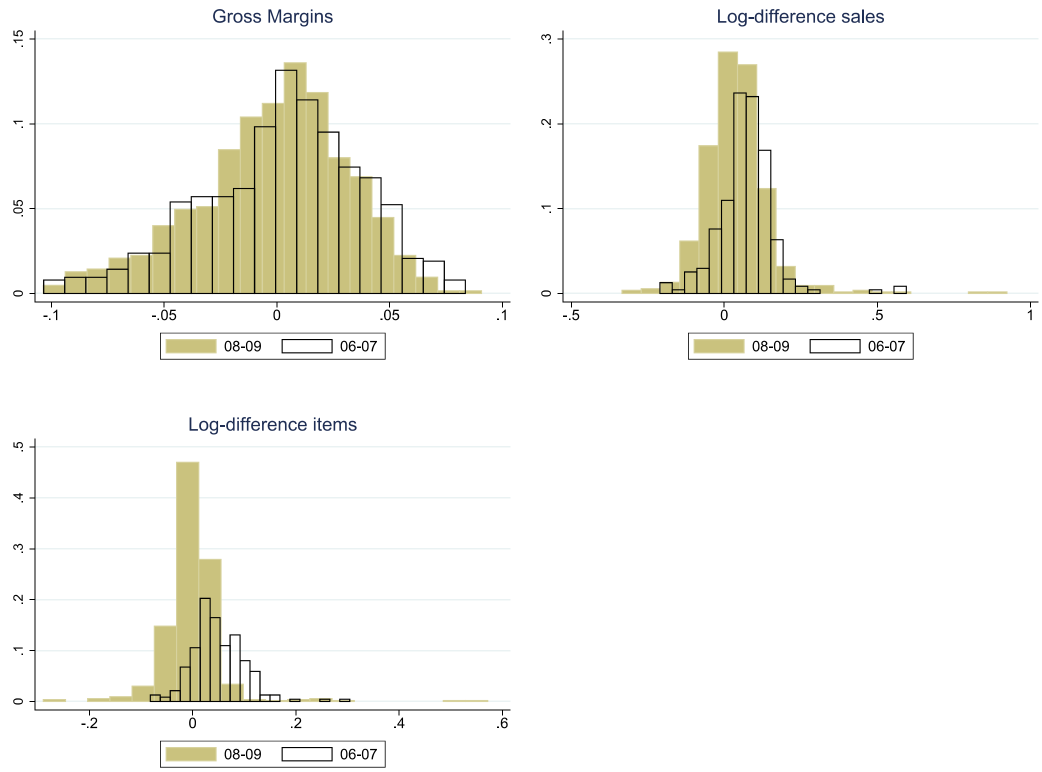
Notes: Variables are log-difference from the prior year. Data is from Compustat and the BLS. The standard deviations are computed at quarterly and annual frequencies, as described in Section 4.2, using data from 1980-2013. We compute the standard deviation of the variable for each firm and then compute the equally-weighted average of the statistic across firms.

Figure 2: Impulse Response Functions to Monetary Policy and Oil Price Shocks



Notes: The figure depicts the impulse response functions of the (log-differenced) gross margins and net operating profit margins to a 1 percentage point monetary policy shock (bottom panel) and an oil price shock (top panel), as described in Section 4.2. The impulse response functions are based on the estimates from the regression equation (4) using quarterly data from Compustat spanning 1985-2009. Standard errors were clustered based on time. The solid lines depict the estimated coefficients and the dashed lines represent the 90th percentile.

Figure 3: Histograms of Gross Margins, Sales, and Number of Items



Notes: Data is from a large U.S. retailer. The figure depicts the distributions of gross margins (levels), sales (log-difference from the prior year), and number of items (log difference from the prior year) for the period 2006-07 and the period 2008-09. Each data point in the distribution observation across regions and time. For confidentiality purposes, we normalize the distribution of gross margin by the mean margin in 2006-07. We do so by subtracting the average 2006-07 margin from the 2006-07 distribution so that the average margin of the normalized distribution is zero. We also subtract the average 2006-07 margin from the 2007-08 distribution. There are 1,256, 771, and 771 observations for gross margins, log difference in sales and log difference in number of items sold, respectively.

Table 5: Cross-sectional Distribution of Margins, Sales, and Number of Items

	Mean	p10	p50	p90
Margins				
Difference	0.007	0.007	0.008	0.007
Log difference in sales				
2006-07	0.025	-0.228	0.029	0.317
2008-09	-0.015	-0.297	0.016	0.236
Difference	-0.040	-0.069	-0.013	-0.082
Log difference in number of items				
2006-07	0.012	-0.134	-0.001	0.158
2008-09	-0.007	-0.150	-0.010	0.145
Difference	-0.019	-0.017	-0.009	-0.013

Notes: Data is from a large U.S. retailer. The table gives key moments from the cross-sectional distribution (across regions) of gross margins, average sales growth, and average growth in the number of items. We report the average levels of each variable in 2006-07 and 2008-09, and the differences between 2006-07 and 2008-09 for sales growth and growth in the number of items. Due to confidentiality reasons, we do not report the *levels* of the margins, and only report how the level of margins changed between 2006-07 and 2008-09.

Table 6: Cyclicality of Store-Item Variables

	Elasticity wrt UR		Elasticity wrt local house prices	
	Coefficient	Standard error	Coefficient	Standard error
Gross margin	-0.003***	(0.001)	-0.003	(0.015)
Price	-0.014***	(0.002)	0.007	(0.019)
Replacement cost	-0.004***	(0.001)	0.001	(0.001)
Sales	-0.066***	(0.013)	0.21***	(0.080)
Number of items	-0.033***	(0.006)	0.151*	(0.086)

Notes: Data is from a large U.S. retailer. Each entry is a separate regression of the log-differenced variable on the local area change in unemployment rate (UR) and house prices, based on regression equation (5) as described in Section 4.4. The regressions with unemployment rates and house prices are based on 2,068 and 58 observations, respectively. Standard errors are clustered by county.

Table 7: Cyclicality of Store-Item Variables: Split by Category

Panel A: US Retailer	Elasticity with respect to UR		Elasticity with respect to house prices	
	Coefficient	SE	Coefficient	SE
Acyclical COGS Categories	-0.004*	(0.002)	-0.132	(0.165)
Procyclical COGS Categories	0.010	(0.009)	0.343	(0.468)
Counter Cyclical COGS Categories	-0.007*	(0.004)	0.103	(0.223)
Panel B: Canadian Retailer	Elasticity with respect to UR		Elasticity with respect to oil prices	
	Coefficient	SE	Coefficient	SE
Acyclical COGS Categories	0.001	(0.003)	-0.093	(0.064)
Procyclical COGS Categories	0.005	(0.026)	-0.082	(0.338)
Counter Cyclical COGS Categories	0.005	(0.003)	-0.054	(0.059)

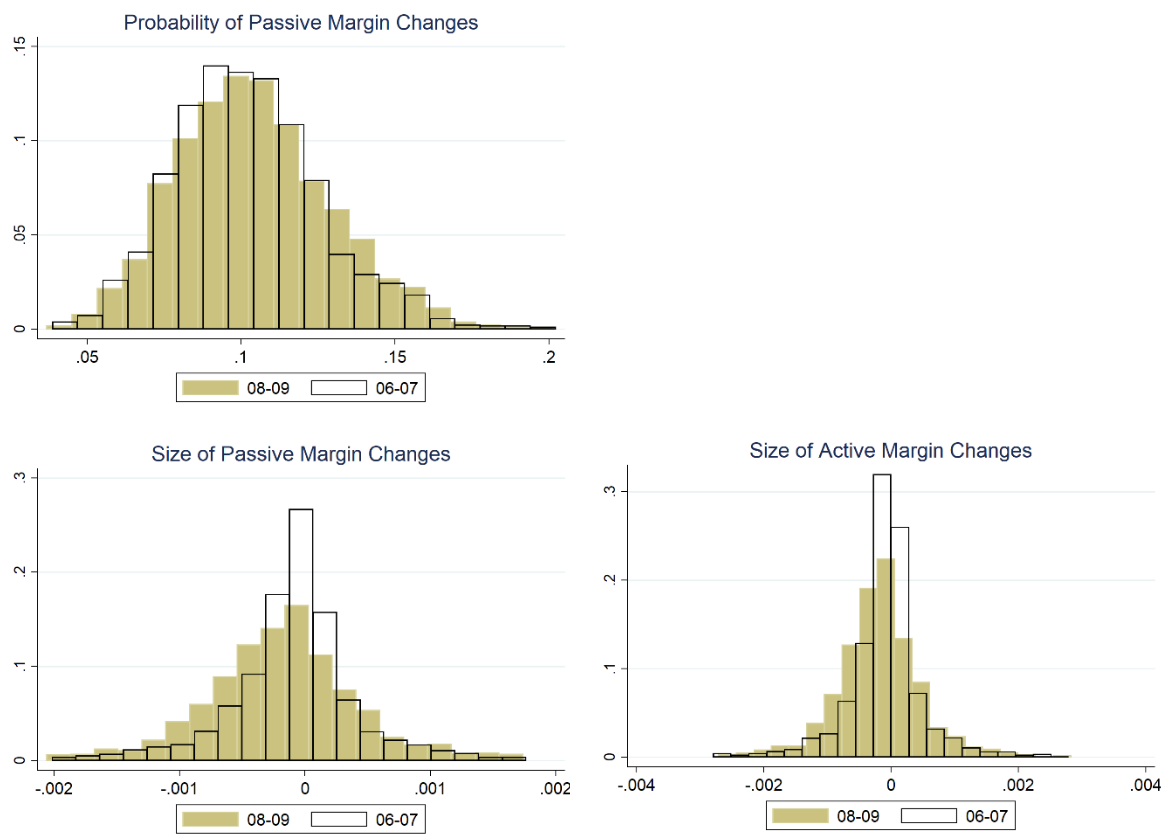
Notes: Panel A uses data from 2006 to 2009 from a large U.S. retailer. Panel B uses data from 2016 to 2018 from a large Canadian retailer. For columns II-IV, each entry is based on a separate regression of the log-differenced gross margins regressed on the local area change in the unemployment rate (UR), based on regression equation 8. For columns V-VII, each entry is a separate regression of the log-differenced gross margins regressed on the local area change in house prices in panel A, based on regression equation 5. For columns V-VII, each entry is a separate regression of the log-differenced gross margins regressed on the local area change in oil prices in panel B, based on regression equation 6. The regressions are run separately for categories that have non-cyclical cost of goods sold (COGS), pro-cyclical COGS, and counter-cyclical COGS. The cyclicality of a category's COGS is based on the category's log-difference replacement costs regressed on the local area change in the unemployment rate. A category is defined as having a non-cyclical COGS if the elasticity of the replacement cost with respect to the unemployment rate is statistically insignificant at a 10 percent level. A category is defined as having a pro-cyclical (counter-cyclical) COGS if the elasticity of the replacement cost with respect to the unemployment rate is negative (positive) and statistically significant at a 10 percent level. For the U.S. retailer, there are 76,448 (1,077) acyclical COGS category observations, 19,999 (276) procyclical COGS category observations, and 20,400 (283) countercyclical COGS category observations for the regressions with unemployment rate (house prices). For the Canadian retailer, there are 30,419 acyclical COGS category observations, 707 procyclical COGS category observations, and 2,543 countercyclical COGS category observations for the unemployment rate and oil price regressions. Standard errors are clustered by region.

Table 8: Active and Passive Margin Changes: U.S. Retailer

	Elasticity wrt UR		Elasticity wrt local house prices	
	Coefficient	Standard error	Coefficient	Standard error
Passive margin changes				
Probability of a passive margin change	0.0008	(0.0015)	-0.013	(0.010)
Size of margin change	0.0000	(0.0001)	-0.002***	(0.001)
Change in replacement cost, given margin change	0.0001	(0.0001)	0.011***	(0.004)
Active margin changes				
Probability of an active margin change	-0.0008	(0.0015)	0.013	(0.010)
Size of margin change	0.0001	(0.0009)	-0.003	(0.008)
Change in replacement cost, given margin change	0.0001	(0.0000)	-0.001	(0.000)

Notes: Data is from a large U.S. retailer, covering the period 2006-2008. Passive margin changes for a given product as those that occur when the replacement cost of that product changes, but the company does not change the product's price. Active margin changes are those that result from changes in the price of that product, independently of whether or not the replacement cost changed. We compute these changes at a daily frequency and then aggregate them at a monthly frequency using the average of the daily changes. Each entry is a separate regression equation 5 for each of the variables, based on 2,021 and 58 observations for the unemployment rate (UR) and house price regressions, respectively. Standard errors are clustered by county. See text for more details.

Figure 4: Histograms of Passive and Active Margin Changes



Notes: Data is from a large U.S. retailer. The figure depicts the probability of passive margin changes, and the distributions of the sizes of passive and active margin changes for the period 2006-07 and the period 2008-09. Each data point in the distribution observation across regions and time. See text for more details.

Table 9: Volatility of Store-Item Variables: U.S. Retailer

Standard Deviation	U.S.
Markup	0.015
Price	0.041
Replacement cost	0.028
Sales	0.220
Number of items	0.115

Notes: Variables are monthly log-difference. Data is from a large U.S. retailer, covering the period 2006-2008. The standard deviations are computed at a monthly frequency. See text for more details.

Table 10: Cyclicality of Store-Item Variables: Canadian Retailer

	Elasticity with respect to local UR		Elasticity with respect to change in oil prices	
Gross margin	0.0001	(0.002)	-0.086	(0.057)
Price	-0.001	(0.005)	0.112*	(0.068)
Replacement cost	-0.001	(0.004)	0.179***	(0.062)
Sales	-0.022**	(0.008)	0.2004	(0.183)
Number of items	-0.015	(0.014)	-0.024	(0.152)

Notes: Data is from a large Canadian retailer, covering the period 2016-2018. Variables are log-differences from prior month. Each row is a separate estimation of regression equation 9, based on 1,267 observations. In columns 1 and 2, the variable are regressed on the local area change in unemployment rate. In columns 3 and 4, each entry gives the estimated coefficient of the differential response of oil producing regions and non-oil producing regions to a change in oil prices. Standard errors are clustered by region. See text for more details.

Table 11: Active and Passive Margin Changes: Canadian Retailer

	Elasticity with respect to local UR		Elasticity with respect to change in oil prices	
Passive margin changes				
Probability of margin change	0.001	(0.003)	-0.123	(0.118)
Size of margin change	0.000	(0.001)	0.002	(0.003)
Change in replacement cost, given margin change	0.002	(0.006)	0.047	(0.148)
Active margin changes				
Probability of margin change	-0.001	(0.003)	0.123	(0.118)
Size of margin change	-0.007**	(0.003)	0.036	(0.057)
Change in replacement cost, given margin change	0.0000	(0.001)	0.007	(0.006)

Notes: Data is from a large Canadian retailer, covering the period 2016-2018. Passive margin changes for a given product as those that occur when the replacement cost of that product changes but the company does not change the product's price. Active margin changes are those that result from changes in the price of that product, independently of whether or not the replacement cost changed. Each row is a separate estimation of regression equation 9, based on 1,267 observations. In columns 1 and 2, the variable are regressed on the local area change in unemployment rate. In columns 3 and 4, each entry gives the estimated coefficient of the differential response of oil producing regions and non-oil producing regions to a change in oil prices. Standard errors are clustered by region. See text for more details.

Table 12: Volatility of Store-Item Variables: Canadian Retailer

	StDev
Gross Margin	0.04
Price	0.06
Replacement cost	0.06
Sales	0.15
Number of items	0.19

Notes: Variables are log-difference from prior month. Data is from a large Canadian retailer, covering the period 2016-2018. The standard deviations are computed at a monthly frequency. See text for more details.

Table 13: Inferring Markups: Various Approaches

Approach:	Output approach	Gross margins approach	Revenue approach
US Retailer:			
Estimated elasticity	0.980	N/A	0.848
	(0.023)	N/A	(0.005)
Markup relative to output approach	1	1.014	0.860
Canadian Retailer:			
Estimated elasticity	0.967	N/A	0.844
	(0.209)	N/A	(0.125)
Markup relative to output approach	1	0.991	0.873

Notes: The first row of each panel in the table reports the estimated output elasticity with respect to the variable input (cost of goods sold). Column 1 reports the elasticity based on the output approach (as described in Bond et al (2020) using the U.S. retailer and Canadian retailer item-level price and replacement cost data over the sample period 2006-2008 and 2016-2018, respectively. Column 3 reports the elasticity based on the revenue approach, which does not use the price and cost data. Column 2 is based on using gross margins as a proxy for markups. This approach does not required any estimated output elasticity. The second row of each panel in the table then reports the inferred markup based on the different approaches. Given confidentiality of the data, we do not report the *level* of the markup. However, we can report how different the inferred markups across the three approaches. Specifically, we report the markup inferred from the gross margin approach and the revenue approach relative to (divided by) the markup inferred from the output approach. See text for more details.

Table 14: Variance Decomposition of the Cross-sectional Margins

	County-level (%) variance	Contribution to total variance
Panel A: U.S. Retailer		
Total	0.117	1.000
Time variation	0.013	0.112
Spatial variation	0.103	0.886
Covariance term	0.000	0.002
Panel B: Canadian Retailer		
Total	0.061	1.000
Time variation	0.017	0.280
Spatial variation	0.051	0.845
Covariance term	-0.008	-0.124

Notes: Data is from a large U.S. retailer over the period 2006-2008 with 3.6 million observations (panel A) and a large Canadian retailer (panel B) over the period 2016-2018 with 15.6 million observations. The table gives the decomposition of the cross-sectional variance (across regions) into four components: differences in gross margins for the same item, differences in assortment composition, the interaction terms, and the covariance terms. See text for more details.

Table 15: Decomposition of the Spatial Variation in Margins

Spatial variation due to:	All items	Item sold everywhere
Panel A: U.S. Retailer		
(i) Differences in gross margins for the same item	14%	10%
(ii) Differences in assortment composition	68%	85%
(iii) Interaction term	1%	1%
(iv) Covariance term	15%	4%
Panel B: Canadian Retailer		
(i) Differences in gross margins for the same item	2%	n.a.
(ii) Differences in assortment composition	61%	n.a.
(iii) Interaction term	3%	n.a.
(iv) Covariance term	34%	n.a.

Notes: Data is from a large U.S. retailer over the period 2006-2008 with 3.6 million observations (panel A) and a large Canadian retailer (panel B) over the period 2016-2018 with 15.6 million observations. The table reports the decomposition of the spatial variance in margins (from page 13, equation 5) into four components. The table reports the average over time of the decomposition. See text for more details.

Table 16: Cross-sectional Variation in Margins and Regional Characteristics

	U.S. Retailer		Canadian Retailer	
	Estimate	Std error	Estimate	Std error
Log household income	0.17***	(0.06)	0.10**	(0.04)
Log median house value	0.16***	(0.05)	0.01	(0.01)
Herfindahl index	-0.01	(0.05)	n.a.	n.a.
Rural county	0.02	(0.01)	n.a.	n.a.

Notes: Data is from a large U.S. retailer containing 80 observations in the cross-section (panel A) and a large Canadian retailer (panel B) with 35 observations in the cross-section. The table reports the elasticity of the gross margin with respect to each of the variables. Each regression is estimated separately. Standard errors are in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Standard errors are clustered by region.

A Online Appendix

Table 17: Cyclicality of Store-Item Variables: U.S. Retailer Margins and Markups

	Elasticity with respect to local UR		Elasticity with respect to local house prices	
Gross margin	-0.003***	(0.001)	-0.0029	(0.015)
Markups	-0.003***	(0.000)	-0.0004	(0.001)

Notes: Data is from a large U.S. retailer, covering the period 2006-2008. Each entry is a separate regression of the log-differenced variable on the local area change in the unemployment rate (UR) and house prices, based on regression equation (8) as described in Section 4.4. The regressions with unemployment rates and house prices are based on 2,068 and 58 observations, respectively. Standard errors are clustered by county. The first row gives the regression based on the log-difference in gross margin, where gross margin is defined as (price-replacement cost)/price. The second row gives the regression based on the log-difference in markup, where markup is defined as price/replacement cost.

Table 18: Cyclicality of Store-Item Variables: Canadian Retailer Margins and Markups

	Elasticity with respect to local UR		Elasticity with respect to change in oil prices	
Gross margin	0.0001	(0.002)	-0.086	(0.057)
Markup	0.0004	(0.002)	-0.040	(0.028)

Notes: Data is from a large Canadian retailer, covering the period 2016-2018. Each entry is a separate regression of the log-differenced variable on the local area change in the unemployment rate and house prices. Each row is a separate estimation of regression equation 9, based on 1,267 observations. Standard errors are clustered by county. The first row gives the regression based on the log-difference in gross margin, where gross margin is defined as (price-replacement cost)/price. The second row gives the regression based on the log-difference in markup, where markup is defined as price/replacement cost. See text for more details.