A Menu Cost Model with Price Experimentation*

David Argente
University of Chicago

Chen Yeh
University of Chicago

February 15, 2016

PRELIMINARY AND INCOMPLETE

Abstract

We document a new set of salient facts on pricing moments over the life-cycle of U.S. products. First, entering products change prices twice as often as the average product. Second, the average size of these adjustments is at least 50 percent larger than the average price change. We argue that a menu cost model with price experimentation can rationalize these findings. The firm is uncertain about its demand elasticity under this setting, but can experiment with its price to endogenously affect its posterior beliefs which are updated in a Bayesian fashion. As a result, firms face the trade-off between increasing the speed of learning through price experimentation and maximizing their static profits. This mechanism can endogenously generate large price changes, without the use of fat-tailed idiosyncratic shocks, and can replicate the life-cycle patterns we document. We show quantitatively that the cumulative output effect of an unanticipated monetary shock is 40 percent larger than in Golosov and Lucas (2007). On impact, selection is weakened as the experimentation motive alters the distribution of desired price changes and decreases the fraction of firms near the margin of adjustment. Furthermore, the notion of a product's life-cycle generates an additional form of cross-sectional heterogeneity in the frequency of price adjustment. This causes the monetary shock to be further propagated.

JEL Classification Numbers: D4, E3, E5

Key words: menu cost, firm learning, optimal control, fixed costs, monetary shocks, hazard rate.

*Contact: `dargente@uchicago.edu` and `chenyeh@uchicago.edu`. We thank Fernando Alvarez, Erik Hurst, Robert Lucas, Robert Shimer and Joseph Vavra for their advice and support. We would also like to thank Bong Geun Choi, Cristián Dagnino, Steve Davis, Jorge García, Elisa Giannone, Mikhail Golosov, Veronica Guerrieri, Munseob Lee, Sara Moreira, Nancy Stokey and seminar participants at the University of Chicago, IEA, EGSC at Washington University, Midwest Macro Meeting, the Bank of Mexico, and EWMES-Milan. David Argente gratefully acknowledges the hospitality of the Bank of Mexico where part of this paper was completed.

1 Introduction

The magnitude of the short-run real effects of monetary policy is an issue that has kept economists debating for decades. After a surge of theoretical frameworks that include a variety of frictions (including sticky prices), the increasing availability of micro-level datasets has allowed us to delve deeper into the mechanics of a firm's dynamic pricing behavior. In recent years, this has lead to a new set of empirical facts that has enhanced our understanding of several moments of the pricing distribution and their interconnections.[1]

Despite new insights of firms' pricing behavior along several dimensions (e.g. different sectors, categories or type of outlets), the degree of price heterogeneity along different stages of the product's *life-cycle* has largely remained unexplored. As a result, this dimension of the data has been almost completely ignored by a broad range of menu cost models. In this paper, we aim to fill this gap by documenting salient facts on the evolution of products' pricing moments over their life-cycle, provide a structural interpretation for it and investigate its implications on monetary non-neutrality in the short run.

It is well known that firms choose different pricing strategies over the life-cycle of their products. As a result, it has been conjectured that firms might have different objectives than merely maximizing current profits.[2] As a result, recognizing and modeling the dynamic patters of prices at different stages of the product life-cycle is crucial to understand the underlying pricing objectives of firms and hence aggregate responses of prices to monetary shocks. This is especially the case since the *nature* of price changes seems to be important for its conclusions on the effectiveness of monetary policy.[3]

In this paper, we use a large panel of barcode level data to document salient facts on a set of pricing moments over the product's life cycle. We calculate the probability of price adjustment (excluding sales) and the size of these adjustments at the weekly level as a function of the product's age. Our main findings are twofold. First, products that enter the market see their prices changed *twice as often* as the average product. Both price increases and price decreases become more frequent whenever demand becomes more uncertain. Second,

[1]Some examples are Golosov and Lucas (2007), Nakamura and Steinsson (2008a), Midrigan (2011), Vavra (2014), and Alvarez and Lippi (2014).

[2]Previously suggested alternatives for these objectives range from survival, market share leadership, liquidating excess inventories upon exit, customer retention and reputation or opposing and eliminating competitive threats through predatory pricing.

[3]An important lesson from the theoretical literature on price-setting is that different types of price changes have substantially different macroeconomic implications. Price changes motivated by a large difference between a firm's current price and its desired price (e.g. Caplin and Spulber (1987)) yield much greater price flexibility than those where the timing of the price change is exogenous (e.g. Calvo (1983)). Similarly, the prevalent view in macroeconomics is that, temporary sales do not play a significant role in inflation dynamics (e.g. Kehoe and Midrigan (2007), Malin et al. (2015)). Recently, however, studies like Kryvtsov and Vincent (2014) have challenged this view.

the average *size* of these adjustments is at least 50% larger than the average price change (i.e. the dispersion of price changes, conditional on adjustment, is higher at entry). Both the frequency and the absolute size of price adjustment approximately settle three months after entry to about 5% and 8% per week respectively.[4] The current class of menu cost models are completely unable to account for these facts as the entirety of pricing moments are independent of the age of the product.

We reconcile the class of menu cost models in the spirit of Golosov and Lucas (2007) with the data by adding price experimentation at the firm-level. When their product is launched, firms are uncertain about their demand. This uncertainty has two roots. Products either belong to a basket of easily substitutable goods or to a group in which it is hard to substitute goods. We assume that a firm is initially unaware to which group its product belongs. Direct inference over the product's life-cycle is then obstructed because of demand shocks. As a result, upon observing a low amount of sales, the firm is unable to distinguish whether this is because they belong to a group with highly substitutable goods or the realization of the demand shock was low.

This uncertainty causes a firm to form beliefs in a Bayesian sense about its type after observing the sold quantity. Our framework features experimentation as a firm is able to alter its posteriors through the endogenously set price. As a result, the firm faces a trade-off between maximizing static profits and deviating in order to affect its posteriors in the most efficient way and acquire more valuable information about its type. At the beginning of its product's launch phase, a firm is completely uncertain about its type and hence has incentives to price experiment in order to gain more information. As the firms learn and obtain sharper posteriors on their type, incentives for experimentation decline, the dispersion of price changes decreases, and the model starts converging to a standard Golosov-Lucas type of model.

We show that the price experimentation motive along with the notion of a product life-cycle allows us to capture two salient features of the data: (1) the life-cycle patterns on the frequency and absolute size of price changes and (2) the presence of large price changes in the data. The latter, in particular, has traditionally eluded standard price-setting models and has been matched by making strong assumptions about the underlying distribution of idiosyncratic shocks which in turn has important implications on monetary non-neutrality.[5]

[4] The patterns at exit are quite different as the frequency and absolute size of price adjustments stay mostly constant before exit at the product level. Nonetheless, the frequency and depth of sales increase significantly at exit suggesting that firms attempt to liquidate their inventory before phasing out their products. These findings are described in more detail in the Appendix.

[5] Midrigan (2011) documents the existence of large price changes and assumes a fat-tailed distribution of cost shocks to match the data. As a result, Midrigan (2011) finds real effects of money in the order of magnitude consistent with those of Calvo-type models. Gertler and Leahy (2008) find a similar result using Poisson shocks. Karadi and Reiff (2012) who use a version of the framework in Midrigan (2011) but instead assume a mixture of two Gaussian distributions as the underlying distribution of idiosyncratic shocks, find

Our framework is able to account for all of these facts simultaneously without having to resort to fat-tailed shocks.

Our framework is most similar to Bachmann and Moscarini (2012). In their model, however, price experimentation is the *only* motive for price changes. As a result, such a framework is not able to account for all the observed regularities in the data. In contrast to their work, our focus is in studying the pricing behavior of firms after they launch a product and our framework nests the standard price-setting model to allow firms to keep changing prices even after firms are almost certain about their type.[6] In addition, the experimentation motives in our framework show strong similarities to those displayed in Keller and Rady (1999) and Mirman et al. (1993). In particular, conditions are developed under which the firm will find optimal to adjust prices away for the myopically optimal level in order to increase the informativeness of the observed market outcomes and as a result increase future profits. We also show that there are two qualitatively different regimes of experimentation, determined by the signal-to-noise ratio, the discount factor, and the initial beliefs of the firm. One regime is characterized by extreme experimentation and a discontinuous optimal policy when the gains of experimentation are larger, the other by moderate experimentation and smaller deviations from the myopically optimal price.

Furthermore, we show that our proposed mechanism implies that the real effects of money are significantly larger than predicted by Golosov and Lucas (2007). The experimentation motive accounts for these results for two main reasons. First, a significant amount of large price changes are purely due to experimentation making them more orthogonal to aggregate money shocks. This reduces the measure of marginal firms whose adjustment decisions are sensitive to nominal shocks, hence weakening the selection effect.[7] And, second, the notion of a products life-cycle generates an additional form of cross-sectional heterogeneity in the frequency of price adjustment. In response to a nominal shock, firms actively learning adjust their price several times before firms with sharper beliefs adjust their price once. Nonetheless, only the first price change matters to accommodate the nominal shock. Given that the model is calibrated to match the average frequency of price changes, the fact that firms who are certain about their type have on average lower frequency of adjustment significantly delays

real effects closer to those found by Golosov and Lucas (2007). These findings indicate that the results on the real effects of money are highly sensitive to the underlying distribution of idiosyncratic shocks.

[6] Bachmann and Moscarini (2012) focus is different as they study how negative first moment shocks induce risky behavior. In their model, when firms observe a string of poor sales, they become pessimistic about their own market power and contemplate exit. At that point, the returns to price experimentation increase as firms “gamble for resurrection”. Our empirical findings indicate however that there is little support for this behavior in the data.

[7] The size-distribution of price changes, under experimentation, has higher kurtosis. Alvarez et al. (2014) show that the kurtosis of the size-distribution of price changes is a sufficient statistic for the real effects of monetary policy.

the adjustment of the aggregate price level propagating the shock even further.

The remainder of the paper is organized as follows. Section 2 presents the data and contains the main empirical findings. In section 3, we set up a quantitative menu cost model that is able to explain the observed facts. In addition, we develop the relevant conditions for experimentation and described the the experimentation regimes in detail. Section 4 discusses our results on monetary non-neutrality and compares our results with other models used in the literature and section 5 concludes.

2 Stylized Facts over Product's Life Cycle

In this section, we use a large scanner data set to show a new set of stylized facts on products over their life-cycle. These facts clearly show that the pricing behavior of firms is considerably different at *entry* compared to the rest of the product's life-cycle. We show that at entry the *frequency* of price adjustment, the *absolute size* of price adjustment and the cross-sectional standard deviation of price changes are higher and settle to their respective averages as the product matures. Furthermore, the fraction of large price changes, defined as those changes larger than two standard deviations, is considerably larger at the beginning of the product life-cycle.

As a result, we hypothesize that these facts can be rationalized with an experimentation motive.[8] The last part of this section provides additional evidence that is consistent with the price experimentation motive by exploiting the variation in products' entry time across different stores.[9]

2.1 Data

The life-cycle patterns of products' prices have typically not been studied much as the requirements on the data are quite stringent. Doing so requires a large panel of products with information about their entry date and their prices at a high sampling frequency. The CPI Research Dataset, for example, is only available at a monthly frequency and contains a left-censored age of the products; many Entry-Level Items (ELIs) are added to the CPI basket long after first appearing in the market. For this reason, we rely on the IRI Marketing data set instead which provides 11 years of weekly store data at the store-week-UPC level. A Universal Product Code (UPC) is a code consisting of 12 numerical digits that is uniquely assigned to a specific tradeable item.

The data set consists of weekly scanner prices and quantity data of drug stores and grocery stores across 50 metropolitan areas (MSA) and 31 product categories.[10] The dataset contains

[8] Optimal control problems in the presence of an experimentation motive, also known as optimal learning problems, have been studied in many areas of economics since Prescott (1972). Its application to the theory of imperfect competition consists of relaxing the assumption that the monopolist knows the demand curve it faces. The first application of this concept can be found in Rothschild (1974). In his paper, the learning process is framed as a two-armed bandit problem in which the monopolist experiments with two prices and eventually settles on one permanently.

[9] Aghion et al. (1991) show that the benefits from experimentation decrease as adequate information is acquired. In Easley and Kiefer (1988) and Kiefer and Nyarko (1989), beliefs also converge to a limit such that learning stops and hence has no additional impact on a firm's decision-making behavior.

[10] These product categories include Beer, Carbonated Beverages, Coffee, Cold Cereal, Deodorant, Diapers, Facial Tissue, Photography Supplies, Frankfurters, Frozen Dinners, Frozen Pizza, Household Cleaners, Cigarettes, Mustard & Ketchup, Mayonnaise, Laundry Detergent, Margarine & Butter, Milk, Paper Towels, Peanut Butter, Razors, Blades, Salty Snacks, Shampoo, Soup, Spaghetti Sauce, Sugar Substitutes, Toilet

around 2.4 billion transactions from January 2001 to December 2011 representing roughly 15% of household spending in the Consumer Expenditure Survey.[11] Each retailer reports the total dollar value of its weekly sales and total units sold for each UPC. Furthermore, it provides information on whether its good was on sale in a certain week (the so-called "sales flag"). This information can then be used to calculate the average retail price in a given week (inclusive of retail features, displays and retailer coupons) by:

$$P_{msctj} = \frac{sales_{msctj}}{units_{msctj}}$$

where m, s, c, t, and j index markets, stores, product categories, time, and UPC respectively. Given the properties of the data, we can identify the first appearance of a UPC in a certain store by exploiting the retail and product identifiers.[12]

2.2 Empirical Strategy

We adopt the same conventions as Coibion et al. (2015) to distinguish between regular price changes and sales. A *regular* price change is defined as any change in prices that is larger than one dollar cent or 1% in absolute value (or more than 0.5% for prices larger than $5 in value) between two periods that neither have a sales flag.

To identify sales, we use the sales flag provided in the data, but our results are robust to applying the sales filter introduced by Nakamura and Steinsson (2008a).[13] The size of a price change is then calculated as the log difference between the price levels in the current

Tissue, Toothbrushes, Toothpaste, and Yogurt. The dataset is discussed in more detail in Bronnenberg et al. (2008). See also Coibion et al. (2015), Gagnon and Lopez-Salido (2014), Chevalier and Kashyap (2014) and Stroebel and Vavra (2014) for applications of the data to macroeconomic questions.

[11] In total, the data covers approximately 170,000 products and around 3,000 stores. Detailed information about each good, such as brand, volume and size, is also included. We exclude private label items from the analysis because all private-label UPCs have the same brand identification so that the identity of the retailer cannot be recovered from the labeling information.

[12] We provide more details of our methodology to identify new product introductions in the Appendix.

[13] Specifically, a good is on sale if a price is reduced but returns to its same previous level within four weeks. Coibion et al. (2015) use two approaches to identify a price spell. The first treats missing values as interrupting price spells. In the second approach, missing values do not interrupt price spells if the price is the same before and after periods of missing values. Since the incidence of sales from applying these two approaches does not significantly differ from the one identified by the sales flag provided in the IRI dataset, we use the union of sales flags obtained from applying these two approaches and the flag provided in the IRI data to identify the incidence of sales. Our results are not sensitive to these choices.

and the previous week.[14] Thus, we have:

$$\Delta P_{msctj} = \log\left(\frac{P_{msctj}}{P_{msct-1j}}\right)$$

Let $a = 1, \ldots, A$ denote the number of weeks since entry (which we will define as the *age* of the product) where $a = 1$ and $a = A$ denote entry and exit respectively. To assess the movements of the pricing moments over the life-cycle of a product, we adopt the following empirical specification:

$$Y_{jstc} = \sum_{a=1}^{A} \phi_a D_{js}^{a} + \theta_{js} + \lambda_t + \gamma_c + \varepsilon_{jstc} \tag{1}$$

where j, s, t and c index the UPC, store, time period, and cohort $c = t - a$ the product belongs to respectively. Y_{jstc} is the variable of interest (i.e. price change indicator, size of price change, frequency of sales); D_{js}^{a} takes the value of one if the product is in its ath week since entry; θ_{js} denotes the fixed effects for each product and store whereas λ_t and γ_c denote time and cohort fixed effects respectively. We are interested in ϕ_a for $a = 1, \ldots, A$ which represent the impact of the life-cycle of the product on our variable of interest.[15]

2.2.1 Pricing Moments over the Products' Life cycle

Figure 1 plots the coefficients of the age dummies for the first 50 weeks of a product using the regular price change indicator as dependent variable. As depicted in the figure, the frequency of price adjustment is almost 4 percentage points higher at entry and takes approximately 20 weeks to settle to its average value. The magnitude of this significantly higher frequency is best reflected in the expected amount of time it takes for a product to change its price.[16] If we were to maintain the frequency of adjustment at entry, then we would expect a price change approximately every 12 weeks. This is *twice* as often relative to the average of 24 weeks that we observe in the data.

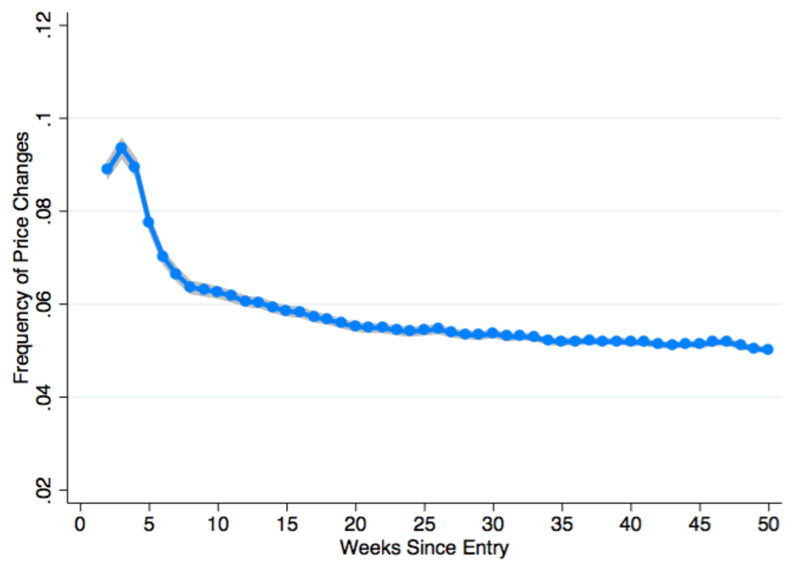
[14] We focus on non-generic products that lasted at least two years in the market and that are sold in at least 20 stores. We do this in order to minimize any type of composition bias that might arise in our calculations from products that last only a few weeks in the market such as seasonal items.

[15] Since it is not possible to identify the effect of age holding time and cohort constant due to collinearity, we estimate equation 1 using two different normalizations. The first assumes that trends appear only in the cohort effects. In this case, we replaced the time fixed effects with the seasonally adjusted unemployment rate at the Metropolitan Statistical Area (MSA) level to control for cyclical economic variation. Some examples of studies using this approach are Deaton and Paxson (1994); De Nardi et al. (2009); Gourinchas and Parker (2002); Aguiar and Hurst (2008). The second assumes trends appear only in the period effects. In this case, only the time fixed effects are included in the estimation of equation 1. Our results are not sensitive to the choice of normalization.

[16] This is equal to $-1/\ln(1-f)$ where f denotes the frequency of price adjustment.

Figure A.4.2 in the appendix decomposes these price changes into increases and decreases. We find that the frequency of both price increases and decreases is higher at entry.

Figure 1: Frequency of Price Adjustment at Entry

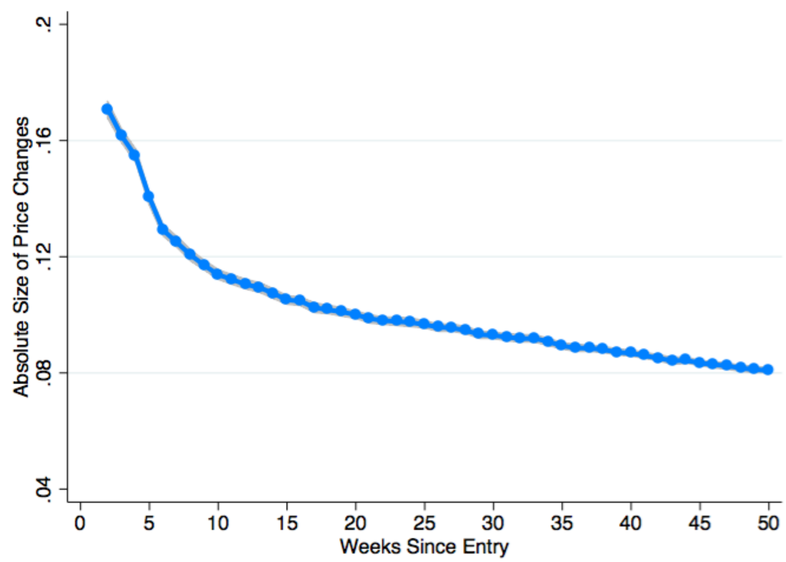


Note: The graph plots the average weekly frequency of price adjustment of entering products. The y-axis denotes the probability that the product adjusts prices in a given week and the x-axis denotes the number of weeks the product has been observed in the data since entry. The graph plots the fixed effects coefficients of equation 1 where we used the regular price change indicator as dependent variable. Equation 1 is computed controlling for store, UPC and time fixed effects. The calculation uses approximately 130 million observations and 2.5 million stores $\times$ UPC pairs. The standard errors are clustered at the store level. The underlying data source is the Symphony IRI.

Similarly, figure 2 depicts the results for the *absolute size* of price changes. During the first few months, the absolute value of price changes is much larger and almost 5 percentage points higher than the average which amounts to approximately 10 percent. Figure A.4.2 in the appendix also shows that in this particular case the variation comes from *both* price increases and decreases. Not surprisingly, the dispersion of price changes, measured by the weekly cross-sectional standard deviation is almost 40 percent larger during the first four months after entry.

Thus, these observations indicate that firms do not only price more often but also in a more extreme fashion in terms of absolute size during the early stages of their products' life-cycles. As a result, we will argue that firm learning is a reasonable mechanism to rationalize our findings in the data. Whenever firms face some form of uncertainty over their demand curve, then these firms are *ex-ante* not able to directly infer the amount of realized sales. As a result, firms form prior beliefs over those demand parameters that are non-observable. This implies that changes in price levels of firms' products can be informative whenever it induces changes

Figure 2: Absolute Value of Price Changes at Entry



Note: The graph plots the average absolute size of price adjustments of entering products. The y-axis is the absolute value of the log price change in that week and the x-axis denotes the number of weeks since the product has entered. The graph plots the fixed effects coefficients of equation 1 where we used the absolute value of the log price change as dependent variable. Equation 1 is computed controlling for store, UPC and time fixed effects. The calculation uses approximately 5.8 million price changes and 2.5 million stores $\times$ UPC pairs. The standard errors are clustered at the store level. The underlying data source is the Symphony IRI.

in their posterior beliefs. Therefore, deviations from profit-maximizing pricing schemes in the form of experimentation can be rational under this particular scenario.

An uncertain monopolist upon entry has learned close to nothing. As a result, its incentives to gather more information and sharpen its posteriors are at its highest during the entering stage of its product(s). Whenever prices can affect posterior beliefs, this monopolist would thus change its prices more often and by larger amounts.

During periods of high levels of experimentation, we should therefore expect to see the frequency of price adjustment and the dispersion of price changes rise. The theoretical implications of these class of models are consistent with the life-cycle patterns described in this section. Furthermore, our findings are consistent with survey evidence by Gaur and Fisher (2005) who found that, by surveying 32 US retailers, 90% of them say they actually conduct price experiments.[17]

[17] An alternative explanation could be penetration pricing. Under this scenario, firms increase long-run profits by launching a low-priced product to secure market share or a solid customer base. This will then result in higher future profits as the firm is able to benefit later from the consumer's higher willingness to pay. In the appendix, we provide evidence that this strategy does not seem to be consistent with our empirical findings. First, we do not find any evidence of lower entry prices. Second, we observe larger price decreases at entry. Lastly and most importantly, the probability of observing consecutive price increases is very low in

2.2.2 Large Price Changes

We now document the patterns of large price changes over the product life-cycle. As those patterns may vary considerably across products and stores, we follow Alvarez et al. (2014) and consider a breakdown of the data into "cells". Each cell is defined by a UPC $\times$ city (MSA) pair. In each cell j, the standardized price change at date t for UPC i is defined as $z_{jit} = (\triangle p_{jit} - \mu_j)/\sigma_j$ where μ_j and σ_j are the mean and standard deviation of price changes in cell j and price changes equal to zero are disregarded.[18]

With this definition in hand and the Symphony IRI data, we are able to document how large price changes are distributed across firms' ages. Equivalently, we can ask ourselves the question of whether large price changes are more or less frequent as firms' products are longer in the market. Figure 3 shows the distribution of regular price changes larger than two standard deviations for the first year and a half since the introduction of that product in the market. The figure shows that, consistent with the empirical prediction of firm experimentation, we observe a sizeable share of large price changes close to entry.[19]

The share of large price changes is considerably larger during roughly the first 10 weeks. About 25% of the price changes larger than two standard deviations occur during these weeks. The distribution of large price changes is roughly uniform after that.[20]

This has the important implication that idiosyncratic shocks from fat-tailed distributions (e.g. Poisson shocks in Midrigan (2011)) cannot be used in menu cost models to account for large price changes. Most of the time, it is assumed that these shocks arrive at a constant rate.[21] As a result, this means that the distribution of large price changes should be independent of age. However, this contradicts our findings in figure 3. Most importantly, the *absence* of fat-tailed shocks has drastic implications on the degree of non-neutrality in the class of menu cost models. This is because the selection effect is stronger whenever the mass of firms on the margin of adjusting is larger. However, this is exactly the case when the distribution of idiosyncratic shocks hitting firms has lower kurtosis (see Alvarez and Lippi (2014)).

2.2.3 The Case for Active Learning and Price Experimentation

The nature of the IRI Symphony data allows us to perform another test that speaks strongly in favor of learning and price experimentation by exploiting the variation in products' entry

our data.

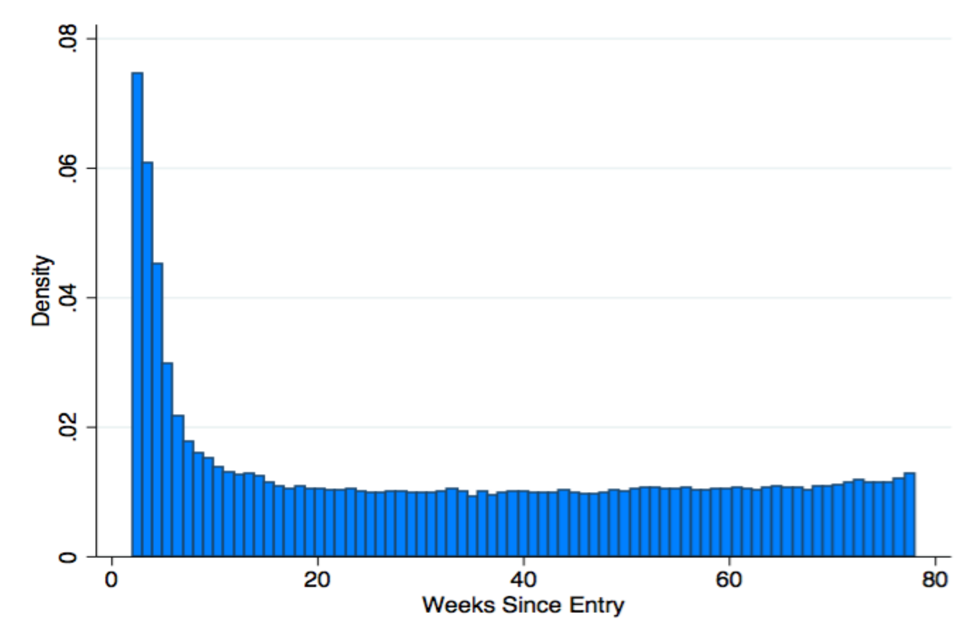
[18]The upper threshold equals $100\times\log(10/3)$ which is about 120 percent log points.

[19]Campbell and Eden (2005) show that extreme prices are relatively young (less than a month old) and argue that grocers deliberately select extreme prices which they quickly abandon in order to experiment.

[20]Figure A.4.3 in the appendix shows that our results are not sensitive to the made standardization since the fraction of non-standardized price changes larger than 30% shows the same pattern.

[21]This includes the family of Poisson shocks used in Midrigan (2011) and Karadi and Reiff (2012).

Figure 3: Fraction of Price Changes Larger than Two Standard Deviations



The figure shows the fraction of price changes larger than 2 standard deviations from the mean in a given category and city as a function of the age of the product. The products considered are those that last at least two years in the market. Source: IRI Symphony dataset

time across space. We observe the introduction of a specific UPC by a retailer across different locations (at the MSA level) and at different times. This allows us to test the hypothesis of whether firms carry forward any information obtained during the *first* launch to any subsequent launch of the same product at a different location. If information on consumer demand is held at the *retailer* level and they are uncertain about some characteristics of this demand curve, then we would expect retailers to experiment less frequently and less aggressively *after* the first launch of their product.[22]

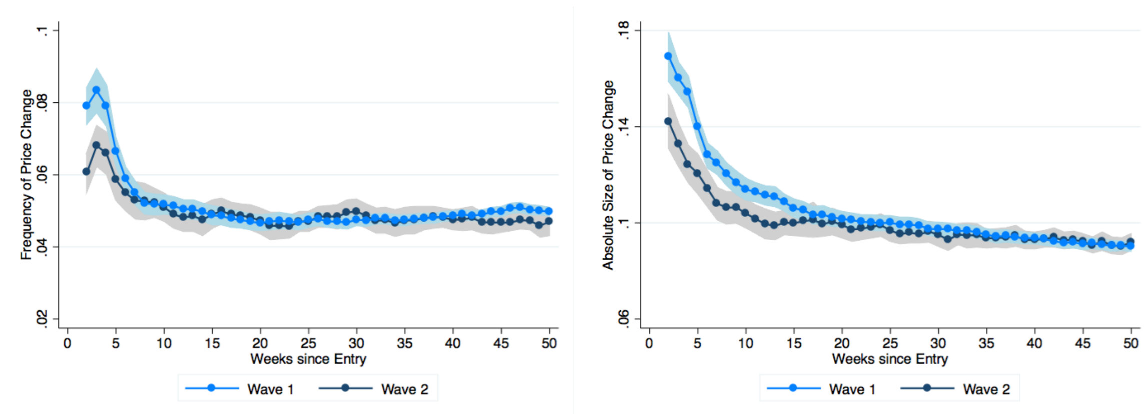
To do this, we first divide every UPC×store pair into two different "waves". A UPC×store pair belongs to the first wave if it was launched by a retailer before the completion of the first year since the UPC was *first* introduced in the market. A UPC×store pair belongs to the second wave if it was introduced by the same retailer at least one year after the product was first launched.

Figure 4 shows that products in the first wave have a higher frequency of adjustment at entry than those in the second wave and shows the same patterns for the absolute size of price changes. On average, at entry, the size of price changes is 7 percent larger than the

[22]Implicitly, we are also assuming that markets across space are not completely independent. This implies, for example, that any information obtained on the local demand of some product in Chicago is at least somewhat informative for another city such as New York.

mean for products in the first wave and only 4 percent larger for products in the second wave. The size of absolute price changes then converges back to the mean of its respective wave.[23] Importantly, these patterns occur for both price increases and decreases. These figures suggests that retailers, after they launch a product, obtain relevant information about the local demand they are facing. When they thus launch the product in the second wave, motives for experimentation are reduced significantly which is exactly reflected in the lower frequency of price adjustment and smaller absolute size of price changes.

Figure 4: Pricing Moments by Waves



(a) Frequency of Price Changes (b) Absolute Size of Price Changes

Note: Panel (a) shows the probability of adjusting prices and panel (b) shows the absolute size of prices changes by waves. Wave 1 represents products that were launched at some location during a period in the first year since the product was introduced *nationally*. Wave 2 represents the same products when launched in different stores a year after its national entry. The graphs shown control for stores, time and product fixed effects.

The remainder of this paper takes these empirical facts as given and assesses the extent to which they can be generated by standard price-setting models. Given that these facts strongly suggest that (active) learning is an important component in firms' pricing behavior, we aim to structurally interpret these facts through a price experimentation mechanism. In the next section, we develop a quantitative menu cost model of a firm that faces demand uncertainty and can obtain more information through its pricing decisions.

[23] In the appendix, we show that these findings also hold if we condition on the fact that waves should occur across different cities.

3 Quantitative Model

Our framework is a discrete time menu cost model in the tradition of Golosov and Lucas (2007) in which firms face uncertainty on their demand curves. As a result, a firm forms beliefs over its type which it can alter through price experimentation. Thus, a firm faces a trade-off between maximizing its static profits and gaining more information about its type when choosing its price. The price experimentation mechanism is based on Mirman et al. (1993) and its implementation in general equilibrium is closely related to Bachmann and Moscarini (2012) (BM henceforth). However, we deviate from BM's framework by removing the firm's fixed costs of production. This eliminates the "gambling for resurrection" effect which we do not observe in our micro-level data. Furthermore, the moments in the data indicate that the frequency and absolute size of price changes converge to a fixed level over time. This indicates that the incentives for price changes are *not* solely driven by price experimentation alone as is the case in BM.[24] We deal with these empirical regularities by adding a menu cost and firm-level idiosyncratic shocks.

3.1 Representative Household

There exists a continuum of differentiated goods consisting of perishable consumption units or services. A good is indexed by a pair (j, k) where the first index $j \in J_1 \cup J_2$ denotes the good's *variety*. Its *type* is denoted by $k \in [0, 1]$.

Goods come in two varieties or baskets: J_1 and J_2 are the groups of *specialties* and *generics* respectively. A group J_i is characterized by its Lebesgue measure γ_i. Varieties within the first basket are hard to substitute with each other whereas generic varieties are mutually substitutable with a relatively high elasticity of substitution. Each good is assumed to be produced by a single monopolistically competitive producer.

There is a unity mass of identical, fully diversified individuals whose preferences for aggregate consumption C and work time H are specified by:

$$\log(C) - \chi \frac{H^{1+\xi}}{1+\xi}$$

Total available time is rescaled to unity, thus H represents the fraction of time spent on working. Consumption C is a Cobb-Douglas aggregate with shares η and $1-\eta$, i.e. we have:

$$C = C_1^{\eta} C_2^{1-\eta}$$

[24] Karadi and Reiff (2012) argue that pricing models which *solely* feature informational or search frictions for price rigidity are difficult to reconcile with moments of the distribution for price changes.

The consumption *sub*aggregates C_1 and C_2 in turn are CES aggregators of monopolistically competitive produced varieties $c_j(k)$.[25] By definition, we get:

$$C_i = \left[\int_{k \in J_i} \alpha_k^{\frac{1}{\sigma_i}} c_i(k)^{\frac{\sigma_i - 1}{\sigma_i}} dk \right]^{\frac{\sigma_i}{\sigma_i - 1}}$$

where $\sigma_2 > \sigma_1 > 1$.[26] Recall that a good is identified through the index pair (j, k). We will assume that j is time-invariant whereas consumer's variety-specific preference shocks α_k are drawn every period, independently over time and across goods. Draws are the same for all consumers. Any consumer takes its consumption and labor decisions *after* observing its taste shocks α_k.

The inverse Frisch elasticity of labor is denoted by ξ and χ represents the relative disutility of labor. Consumers elastically supply labor H on a competitive market at wage rate W. Furthermore, it owns a fully diversified portfolio consisting of assets of all the firms in the economy. Thus, it receives profits Π from firm ownership.

It is assumed that households have *no savings device*, thus we completely abstract from the intertemporal problem.[27] As a result, the consumer's problem is static. We can break the consumer problem into a sequential one with two stages. In the first stage, the consumer picks the amount of expenditure allocated to subaggregates C_i and its labor supply decision H. Let us suppose there exist price indices $\overline{P}_i$ such that $\overline{P}_1 C_1 + \overline{P}_2 C_2 = S$ where S denotes total spending. We will verify later what these indices exactly look like. Then, the problem reduces to:

$$\max_{H, C_1, C_2} \log\left(C_1^{\eta} C_2^{1-\eta}\right) - \chi \frac{H^{1+\xi}}{1+\xi} \quad \text{s.t.} \quad \overline{P}_1 C_1 + \overline{P}_2 C_2 = S. \qquad (\nu)$$

where ν denotes the Lagrangrian multiplier corresponding to the budget constraint. Standard Cobb-Douglas logic implies that a fraction η and $1-\eta$ of total income will be spent on baskets 1 and 2 respectively. Thus, we get:

$$C_1 = \frac{\eta S}{\overline{P}_1} \quad \text{and} \quad C_2 = \frac{(1-\eta) S}{\overline{P}_2}.$$

[25] Note that we deviate from BM's notion of the consumption aggregate in which $C_i^{\sigma_i - 1/\sigma_i} = \int_{J_i} \int_0^1 \alpha_k^{1/\sigma_i} c_j(k)^{\sigma_i - 1/\sigma_i} dk dj$. Under this interpretation, a consumption basket consists of a continuum of continuums. In our context of price experimentation, this is without loss of generality. Our results on firm experimentation are unaffected as long as the consumer's demand function is separable in a log-linear fashion.

[26] This explains our terminology for the basket groups J_1 and J_2. We assume that $\sigma_1 < \sigma_2$, hence goods within J_1 are harder to substitute. Furthermore, we assume that, *within a variety*, goods are spread out *equally*. Formally, this is denoted by $k \sim U[0, 1]$.

[27] Alternatively, we could have assumed that consumers have completely restricted access to capital markets.

The first order condition with respect to H immediately implies:

$$\chi H^{\xi} = \nu W$$

Our Cobb-Douglas specification also tells us that the marginal utility of income is simply the inverse of total spending, thus we must have $\nu = S^{-1}$. By assumption, total spending is simply the sum of labor income and profits due to firm ownership, i.e. $S = WH + \Pi$. As a result, we obtain:

$$H^{\xi} = \frac{W}{\chi} \frac{1}{WH + \Pi}$$

The implicit solution to the above equation defines the labor supply decision $H^s(W, \Pi)$. Whenever W, χ and Π are strictly positive, then $H^s(W, \Pi)$ is guaranteed to exist.

The previous derivation showed us that a fraction η and $1-\eta$ of the consumer's income S is spent on each basket. Thus, the second stage problem immediately gives us the standard, downward-sloping CES demand curve.

$$c_i(k) = \alpha_k \left(\frac{P_i(k)}{\overline{P}_i} \right)^{-\sigma_i} \frac{\eta_i S}{\overline{P}_i} \tag{2}$$

where $\eta_1 = \eta$ and $\eta_2 = 1 - \eta$ with some abuse of notation. This result follows whenever we define the aggregate price index $\overline{P}_i$ as:

$$\overline{P}_i = \left(\int_{k \in J_i} \alpha_k P_i(k)^{1-\sigma_i} dk \right)^{1/1-\sigma_i} .$$

Note that $\overline{P}_i$ is *independent* of the *actual* realizations of the demand shocks. Since there is a continuum of goods within each basket i, we can exploit a law of large numbers.[28] This implies:

$$\overline{P}_i = \left(\int_{k \in J_i} P_i(k)^{1-\sigma_i} dk \right)^{1/1-\sigma_i}$$

where our law of large numbers gives us that $\int_0^1 \alpha_k dk \xrightarrow{p} \mathbb{E}_\alpha(\alpha_k)$. Then, the previously equality follows from the normalization $\mathbb{E}_\alpha(\alpha_k) = 1$. Without loss of generality, we impose the structure $\alpha_k = \exp(F^{-1}(k))$ on the preference shocks which will prove to be useful later.

[28] In particular, we exploit the *Glivenko-Cantelli* theorem. We employ a similar argument as BM.

As a result, we can write:

$$\begin{aligned}\mathbb{P}(\varepsilon_k \leq \epsilon) &\equiv \mathbb{P}(\log(\alpha_k) \leq \varepsilon)\\ &= \mathbb{P}(F^{-1}(k) \leq \varepsilon)\\ &= \mathbb{P}(k \leq F(\varepsilon))\\ &= F(\varepsilon)\end{aligned}$$

where the last equality follows from the fact that k is uniform on J_i for each $i \in \{1, 2\}$.

3.2 Firms

Firms are characterized by a pair (j, k) and produce in a monopolistically competitive fashion. Every firm takes consumer demand as given, but does *not* know whether its variety is part of the specialty or generics basket, i.e. it does not know whether $j = 1$ or $j = 2$. It also cannot observe the realization of the taste shock α_k even though it does observe the *realized* demand quantity after setting its prices optimally. Intuitively, this is the root of the informational problem: a firm does not know whether consumer demand is high due to setting low prices and belonging to the elastic basket (i.e. $\sigma_i = \sigma_2 > \sigma_1$) or whether taste shocks for its variety are simply high. However, realized sales are observable and most importantly informative. As a result, a firm can update its beliefs in a Bayesian fashion. The firms' production technology consists solely of labor and is linear. Lastly, we assume that a firm can exit due to exogenous reasons. This is given by the death rate δ.[29]

3.2.1 Information Structure

In contrast to most state-dependent pricing models, firms in our framework set prices under incomplete information. Any firm (j, k) with $j \in J_i$ is unaware of the realization of α_k and σ_i. As a result, a firm might want to deviate from the static profit maximizing price and experiment to learn more about the price elasticity of its corresponding basket. At any point

[29]The death rate faced by firms is due to *force majeure* reasons unrelated to profitability. A plethora of frameworks work with this assumption at either the product- or firm-level (Bernard et al. (2007) and Melitz (2003) respectively). Alternatively, we could have ignored the extensive margin at the product-level and used informational Poisson shocks as in Baley and Blanco (2014). However, we refrain from this for mainly two reasons. First, the IRI data speaks clearly on product-level entry and exit and its importance on pricing moments. These characteristics of the pricing distribution is exactly what we are trying to capture. Second, idiosyncratic shocks from fat-tailed distributions are known to have considerable real effects on monetary non-neutrality (e.g. Midrigan (2011) and Karadi and Reiff (2012)). Our goal is to assess the effect of firm-level price experimentation on the effectiveness of monetary policy. We believe our results are more transparent without the interaction of fat-tailed shocks.

in time, a firm can observe its amount of sales $Q_j(k)$ after setting some price $P_j(k)$.[30] For any variety belonging to basket J_i, we obtain:

$$Q_j(k) = \alpha_k \left(\frac{P_j(k)}{\overline{P}_i} \right)^{-\sigma_i} \frac{\eta_i S}{\overline{P}_i}$$

Obviously, this demand specification is log-linearly separable. We obtain:

$$\log\left(Q_j(k)\right) = \log(\alpha_k) - \sigma_i \log\left(\frac{P_j(k)}{\overline{P}_i}\right) + \log(\eta_i S) - \log(\overline{P}_i)$$

Let non-capitalized letters denote its logged counterpart, then we get:

$$\begin{aligned} q_j(k) &= -\sigma_i p_j(k) + s + (\sigma_i - 1)\overline{p}_i + \log(\eta_i) + \log(\alpha_k) \\ &= -\sigma_i p_j(k) + s + \mu_i + \varepsilon_k \end{aligned}$$

where we defined $\varepsilon_k = \log(\alpha_k)$ and $\mu_i = (\sigma_i - 1)\overline{p}_i + \log(\eta_i)$. Note that this specification does *not* result into a one-to-one mapping between quantities and prices. Whenever a firm sets a price $p_j(k) \equiv \log(P_j(k))$, demand $q_j(k)$ can be high for three reasons: (1) the variety belongs to a basket within which substitution is hard (and therefore market power is high), i.e. $\sigma_i = \sigma_1$, (2) consumers have a strong preference for variety (j, k) due to a high realization of ε_k or (3) the variety to which j belongs to has only a few competitors (i.e. low γ_i) so $\overline{p}_i \equiv \log(\overline{P}_i)$ is high and the consumer spends much of her income on goods in basket i.[31]

Our quantitative results depend on CES preferences which are used in most of the price-setting models. However, our results on firm-level price experimentation do not rely on these type of preferences alone. As long as the demand function is linearly separable after some uniform transformation, then the results hold.[32]

[30] Note that we directly imposed goods market clearing at the variety level, i.e. $Q_j(k) = c_j(k)$ for any pair (j, k). We will assume that a firm is contractually obliged to deliver the ex-post quantity of demanded goods to the consumer. This will be made explicit in section 3.2.2. However, a firm sets its price *ex-ante* before the realization of the demand shocks.

[31] To see this in the most transparent way, consider the following simplification. Imagine that a firm knows its elasticity of substitution σ_i but does not know the realization of the taste shock α_k and hence takes expectations over it. As a result, the optimal set price is characterized by $P_j^*(k) = \frac{\sigma_i}{\sigma_i - 1}\frac{w}{z}$. Then, we immediately can deduce $\overline{P}_i = \left(\frac{\sigma_i}{\sigma_i-1}\frac{w}{z}\right)\gamma_i^{1/1-\sigma_i}$ and $\overline{p}_i = \frac{1}{1-\sigma_i}\log(\gamma_i) + \log\left(\frac{\sigma_i}{\sigma_i-1}\frac{w}{z}\right)$. Since $\sigma_i > 1$, it is clear that $\overline{p}_i$ is strictly decreasing in γ_i.

[32] A subset of frameworks on price experimentation use the linear demand specification directly. Keller and Rady (1999) demonstrate the potential power of a linear demand curve as they obtain analytical expressions on a firm's pricing policy function under incomplete information. However, it is unclear whether this form of experimentation is reconcilable with the empirical facts found in section 2. Willems (2013) adopts a linear demand curve in a setting that is akin to ours. However, his setup is unsuitable for the analysis of the effectiveness of monetary policy as it is not closed in general equilibrium. Furthermore, firms engage in active learning (i.e. experimentation) under the presence of good-specific habits as individual price volatility is mainly rooted between the interaction of the active learning motive with the presence of these habits.

3.2.2 Bayesian Updating

The pair that uniquely identifies a firm is not observed by the firm itself. As a result, its pricing policy is independent of (j, k) and we can drop this indexing without loss of generality. Our setup imposes the following timing on the firm's pricing decisions and the consumer's realized demand shock for each period.

1. A firm decides on a price $p = \log(P)$ *before* the realization of the log-demand shock ε_k.

2. The shock ε_k is realized and the representative consumer decides how much to consume.

3. The firm is contractually obliged to supply $\exp(q) = \exp(-\sigma_i p + s + \mu_i + \varepsilon_k)$.

Even though the firm can only observe q, the realized quantity sold or its sales are *informative*. Given some prior belief λ of being a $\sigma = \sigma_1$ type firm, it can update this prior to some posterior λ'. Let F_i denote the distribution of ε whenever the firm's true type is $\sigma = \sigma_i$, then a simple application of Bayes' rule gives us:

$$\begin{aligned}\lambda' &= B(\lambda, p, q, s)\\ &= \frac{\lambda F_1'(q + \sigma_1 p - \mu_1 - s)}{\lambda F_1'(q + \sigma_1 p - \mu_1 - s) + (1-\lambda) F_2'(q + \sigma_2 p - \mu_2 - s)}\\ &= \left[1 + \frac{1-\lambda}{\lambda}\frac{F_2'(q + \sigma_2 p - \mu_2 - s)}{F_1'(q + \sigma_1 p - \mu_1 - s)}\right]^{-1}\end{aligned}$$

However, dynamic decision-making requires knowing the evolution of beliefs *conditional* on a given state i, where true sales are $q = -\sigma_i p + s + \mu_i + \varepsilon$. The firm rationally anticipates that a price change will affect the informative quantity it will observe directly after. A price change in the current period affects a firm's *future* beliefs about being in basket 1 conditional on the true state being i. As a result, the firm's motives are not solely rooted in the maximization of its static profits as a firm's pricing strategy can increase the value of its sales' informativeness.

To illustrate our point as clearly as possible, suppose the true state is $i = 1$. This implies that the "sales data" must be generated by $q = -\sigma_1 p + s + \mu_1 + \varepsilon$ which is what the firm does observe. A firm is uncertain about whether its product belongs to basket 1 or 2 though and cannot observe the log demand shock ε. Suppose tomorrow's realized taste shock is ε_j whenever a firm believes the state to be j.

Unfortunately, this generates price dynamics that are not consistent with our empirical observations.

Then, for a prior λ, Bayesian updating gives us:

$$\begin{aligned}
b_1(\lambda, p, \varepsilon) &= \frac{\lambda F'(\varepsilon)}{\lambda F(\varepsilon) + (1-\lambda)F'(\mu_1 - \mu_2 + (\sigma_2 - \sigma_1)p + \varepsilon)} \\
&= \left[1 + \frac{1-\lambda}{\lambda}\frac{F'(\mu_1 - \mu_2 + (\sigma_2 - \sigma_1)p + \varepsilon)}{F(\varepsilon)}\right]^{-1} \qquad (3)
\end{aligned}$$

The term $F(\varepsilon)$ is straightforward: whenever a firm beliefs to be in state 1 then it thinks that $\sigma = \sigma_1$ and $\mu = \mu_1$. As a result, the shock ε_1 only needs to be equal to ε to be consistent with *observed* log sales q. The last probability term $F'(\mu_1 - \mu_2 + (\sigma_2 - \sigma_1)p + \varepsilon)$ requires some explanation. Whenever a firm beliefs that the state is 2 then it presumes $\sigma = \sigma_2$ and $\mu = \mu_2$. However, it still observes q which was generated through $q = -\sigma_1 p + s + \mu_1 + \varepsilon$ (as the true state equals 1). Thus, the shock ε_2 needs be such that it is consistent with the true data generating process as conjectured by the firm. Thus, we require:

$$\begin{aligned}
-\sigma_2 p + s + \mu_2 + \varepsilon_2 &= -\sigma_1 p + s + \mu_1 + \varepsilon \\
&= q
\end{aligned}$$

Rearranging for ε_2 from the first equality, we get $\varepsilon_2 = \mu_1 - \mu_2 + (\sigma_2 - \sigma_1)p + \varepsilon$. This occurs with probability $F'(\varepsilon_2) = F'(\mu_1 - \mu_2 + (\sigma_2 - \sigma_1)p + \varepsilon)$ which explains expression (2). A similar argument can be imposed whenever the true state is $i = 2$. The Bayesian updating formula (for believing to be in basket 1) only changes slightly to:

$$\begin{aligned}
b_2(\lambda, p, \varepsilon) &= \frac{\lambda F'(\mu_2 - \mu_1 + (\sigma_1 - \sigma_2)p + \varepsilon)}{\lambda F'(\mu_2 - \mu_1 + (\sigma_1 - \sigma_2)p + \varepsilon) + (1-\lambda)F'(\varepsilon)} \\
&= \left[1 + \frac{1-\lambda}{\lambda}\frac{F'(\varepsilon)}{F'(\mu_2 - \mu_1 + (\sigma_1 - \sigma_2)p + \varepsilon)}\right]^{-1} \qquad (4)
\end{aligned}$$

The updating formulas (3) and (4) can be compressed into:

$$b_i(\lambda, p, \varepsilon) = \left[1 + \frac{1-\lambda}{\lambda}\frac{F'(\mu_i - \mu_2 + (\sigma_2 - \sigma_i)p + \varepsilon)}{F'(\mu_i - \mu_1 + (\sigma_1 - \sigma_i)p + \varepsilon)}\right]^{-1} \qquad (5)$$

which is exactly equal to the posterior B evaluated at the true data generating process for sales. Thus, we obtain:

$$b_i(\lambda, p, \varepsilon) = B(\lambda, p, -\sigma_i p + s + \mu_i + \varepsilon, s).$$

3.2.3 Speed of Learning

The speed of learning or the sensitivity of any price change on a firm's posterior belief is completely captured by the blue term in expression (4):

$$\frac{F'(\mu_i - \mu_2 + (\sigma_2 - \sigma_i)p + \varepsilon)}{F'(\mu_i - \mu_1 + (\sigma_1 - \sigma_i)p + \varepsilon)}$$

which is the *likelihood ratio* under the null of state i being the true one. Recall $\mu_i = (\sigma_i - 1)\overline{p}_i + \log(\eta_i)$ by definition and suppose $i = 1$, then we can also write the likelihood ratio as:

$$\begin{aligned}\frac{F'(\mu_1 - \mu_2 + (\sigma_2 - \sigma_1)p + \varepsilon)}{F'(\varepsilon)} &= \frac{F'((\sigma_1 - 1)(\overline{p}_1 - p) - (\sigma_2 - 1)(\overline{p}_2 - p) + \varepsilon)}{F'(\varepsilon)} \\ &= \frac{F'\left((\sigma_1 - 1)\overline{p}_1 - (\sigma_2 - 1)\overline{p}_2 + \varepsilon\left[1 + \frac{\sigma_2 - \sigma_1}{\varepsilon}p\right]\right)}{F'(\varepsilon)}\end{aligned}$$

Intuitively, the price indices $\overline{p}_i$ harbor information on a firm's competitors as it reflects the basket's "average" price. Therefore, a deviation from these averages make sales more informative as a firm gains additional informative signals on top of the information contained in the aggregate prices. Note that we only care about the *absolute* price deviation $|\overline{p}_i - p|$, thus large deviations from the aggregate price indices, be it negative or positive, both provide much information.

Lastly, note that the impact of prices on posteriors are larger (and hence, more incentives for experimentation) whenever the signal-to-noise ratio $(\sigma_2 - \sigma_1)/\sigma_\varepsilon \equiv \Delta\sigma/\sigma_\varepsilon$ is high. Whenever $\varepsilon = \log(\alpha) \sim N(m, \sigma_\varepsilon^2)$, we can derive that:

$$b_i(\lambda, p, \xi) = \left(1 + \frac{1-\lambda}{\lambda}\exp\left(\tfrac{1}{2}(-1)^{\mathbf{1}(i=2)}\left[\left(\frac{\xi}{\sigma_\varepsilon}\right)^2 - \left(p\frac{\Delta\sigma}{\sigma_\varepsilon} - \frac{\Delta\mu}{\sigma_\varepsilon} + \frac{\xi}{\sigma_\varepsilon}\right)^2\right]\right)\right)^{-1}$$

The above expression then clearly indicates that firms get more *bang-for-buck* whenever the signal-to-noise ratio is high. We will choose a baseline specification for our quantitative results in which $m = 0$ and σ_ε is inferred from the data.

3.2.4 Production Technologies

Firms are *ex-ante* identical but can generate heterogeneous *ex-post* pricing paths as different realizations of the log demand shocks induce differently updated posterior beliefs. All firms have access to a linear production technology in which $1/z$ units of labor can be transformed into one unit of output, i.e. we have:

$$y_{j,k}(\ell) = z\ell$$

As a result, its marginal costs of production are constant and equal to w/z. Thus, a firm's *static* profits equal:

$$\begin{aligned}\widetilde{\Pi}_i(P;\alpha_k) &= \left(P - \frac{w}{z}\right) Q(P;\alpha_k) \\ &= \left(P - \frac{w}{z}\right) \alpha_k \left(\tfrac{P}{\overline{P}_i}\right)^{-\sigma_i} \tfrac{\eta_i S}{\overline{P}_i}\end{aligned}$$

3.2.5 Firm's Dynamic Pricing Policies under Incomplete Information

A firm chooses a price to determine the trade-off between maximizing current profits and obtaining more accurate information in the future about its elasticity of demand. Since a firm cannot observe the realization of the demand shock α_k whenever it has to decide on its pricing policy, it has to take expectations over α_k. Due to our normalization $\mathbb{E}_\alpha(\alpha_k) = 1$, we obtain a firm's *ex-interim* expected profits:

$$\begin{aligned}\Pi_i(P) &\equiv \mathbb{E}_\alpha\left[\widetilde{\Pi}_i(P;\alpha_k)\right] \\ &= \left(P - \frac{w}{z}\right)\left(\frac{P}{\overline{P}_i}\right)^{-\sigma_i} \frac{\eta_i S}{\overline{P}_i}\end{aligned}$$

Any firm prices while taking aggregate prices, spending and the wage rate as given.[33] These variables are determined in general equilibrium and summarized by the aggregate state $\omega \equiv (\overline{P}_1, \overline{P}_2, W, S) \in \Omega$. Its Markov transition density will be denoted by $\mathcal{T}(\omega';\omega)$. This density can be explicitly derived through the laws of motions for the aggregate price indices $\overline{P}_i$. Every firm also carries a belief λ about being a type 1 firm. Furthermore, it discounts at the rate $\beta \in (0,1)$.[34]

[33] In a world without a need for learning, i.e. the firms knows to which basket J_i it belongs, the firm faces a static problem every period and would ideally like to set its price to maximize its static profits *given* some fixed productivity z. This leads to the standard CES mark-up over marginal costs that is observed in many frameworks as $P_i^*(z) = \frac{\sigma_i}{\sigma_i - 1}\frac{W}{z}$. This is useful intuitively as price experimentation is bounded in this interval, i.e. we should expect:

$$P^*(\lambda, z) \in [P_2^*(z), P_1^*(z)]$$

Recall that $\sigma_1 < \sigma_2$, thus firms belonging to basket 1 face less elastic demand schedules. As a result, they have more market power and should set higher mark-ups. The degree of price variability that we observe in the data can be used to calibrate the parameters σ_1 and σ_2 as these two parameters govern the bounds P_2^* and P_1^*.

[34] The model implicitly assumes that any firm dies with some exogenous probability δ. It is assumed that

Furthermore, the firm incurs a menu cost when it decides to change its price: when the firm endogenously decides on a price $P_t \neq P_{t-1}$, it has to incur a menu cost of ψ which is denoted in terms of labor units. Lastly, we incorporate idiosyncratic productivity shocks at the firm-level. We assume that the idiosyncratic productivity term z_t follows an AR(1) process:

$$\log(z_{t+1}) = \rho \log(z_t) + \sigma_\zeta \zeta_{t+1} \text{ where } \zeta_{t+1} \sim N(0, 1)$$

Thus, our framework is a hybrid version of standard state-depending pricing models and frameworks featuring price experimentation.

As a result, our framework differs from BM in two fundamental ways. First, our framework features a menu cost rather than a fixed cost of production. This will deliver two features consistent with our observations of section 2: (1) without a fixed cost of production, our framework does not feature the "gambling for resurrection" effect as found in BM. We choose to do this as we do not find any evidence for this effect in our data, (2) a *menu* cost will match the empirical frequencies of price changes. This is not featured in the baseline case of BM in which a firm will eventually learn its type over time but changes its prices every period in reaction to even the slightest change in observed sales as it is costless to do so.

Second, the incentives for price changes at the later stages of a product's life-cycle seem to be better captured by idiosyncrasies. Without idiosyncratic shocks but with a menu cost present, a firm would have no incentives to change its price as at the late stage of its product's life-cycle, most information about its type is already harvested due to earlier price experimentation. As a result, the gains to obtaining additional information at this stage are extremely small and never offset by a menu cost. However, we observe that the frequency and absolute size of price changes at these stages are non-negligible. Thus, we capture these price changes through standard state-contingent channels by incorporating idiosyncratic cost shocks and allowing for positive inflation levels.

To allow for inflation, we assume that nominal spending $S_t = \overline{P}_t C_t$ grows at a constant rate $\tilde{\pi} \geq 0$:

$$\log(S_{t+1}) = \log(S_t) + \tilde{\pi}$$

We will focus on an equilibrium in which real variables stay constant. This implies that the aggregate state satisfies $\omega_t = \omega^*$ for some $\omega^* \in \Omega$ and the transition density collapses to a mass point at ω^*. As a result, a firm's dynamic programming problem is summarized by the

this exogenous death rate is already incorporated in the discount factor β.

following Bellman equation:

$$V(\lambda, z, p_{-1}) = \max\left\{V^A(\lambda, z), V^N(\lambda, z, p_{-1})\right\}$$

$$\begin{aligned}
\text{where} \quad V^A(\lambda, z) &= \max_{p \geq 0} \left(p - \tfrac{W}{z}\right) \left[\lambda\eta \tfrac{p^{-\sigma_1}}{P_1^{1-\sigma_1}} + (1-\lambda)(1-\eta) \tfrac{p^{-\sigma_2}}{P_2^{1-\sigma_2}}\right] \tfrac{S}{P} - \psi \tfrac{W}{P} \\
&+ \beta\lambda \int_{z'} \int_{\epsilon} V\left(b_1(\lambda, \log(\tfrac{p}{1+\tilde{\pi}}), \epsilon\right), z', \tfrac{p}{1+\tilde{\pi}}) dF(\epsilon) dG(z', z) \\
&+ \beta(1-\lambda) \int_{z'} \int_{\epsilon} V\left(b_2(\lambda, \log(\tfrac{p}{1+\tilde{\pi}}), \epsilon\right), z', \tfrac{p}{1+\tilde{\pi}}) dF(\epsilon) dG(z', z) \\
V^N(\lambda, z, p_{-1}) &= \left(p_{-1} - \tfrac{W}{z}\right) \left[\lambda\eta \tfrac{p_{-1}^{-\sigma_1}}{P_1^{1-\sigma_1}} + (1-\lambda)(1-\eta) \tfrac{p_{-1}^{-\sigma_2}}{P_2^{1-\sigma_2}}\right] \tfrac{S}{P} \\
&+ \beta\lambda \int_{z'} \int_{\epsilon} V\left(b_1(\lambda, \log(\tfrac{p_{-1}}{1+\tilde{\pi}}), \epsilon\right), z', \tfrac{p_{-1}}{1+\tilde{\pi}}) dF(\epsilon) dG(z', z) \\
&+ \beta(1-\lambda) \int_{z'} \int_{\epsilon} V\left(b_2(\lambda, \log(\tfrac{p_{-1}}{1+\tilde{\pi}}), \epsilon\right), z', \tfrac{p_{-1}}{1+\tilde{\pi}}) dF(\epsilon) dG(z', z)
\end{aligned}$$

Recall that a firm does not observe the taste shocks for its specific variety and is not aware to which basket its good belongs to. As a result, the optimally chosen price must be independent of k and j. We will define the optimal pricing policy $P^*(\lambda, z)$ as the maximizer associated with the value function $V^A(\lambda, z)$.

In a menu cost model *without* experimentation, a price-setting firm would only consider its static profits and its effect on the continuation value through the price level tomorrow as changing its price is not costless.[35] However, sales are observable and informative. Thus, a firm can affect its posterior beliefs through its price. This is highlighted by the terms b_1 and b_2 in the firm's continuation value. As a result, the policy function $P^*(\lambda, z)$ reflects the optimal deviation from the myopic policy function which summarizes the balance between sacrificing static profits and sharpening its posteriors beliefs.[36]

[35]This class of frameworks include standard price-setting models such as Barro (1972), Dixit (1991), Golosov and Lucas (2007) and Alvarez and Lippi (2014) for example.

[36]Note that our framework is fundamentally different from most price-setting models with *learning*. In the framework by Baley and Blanco (2014), a firm is faced with uncertainty about its productivity. As a result, the problem can be described as a Kalman-Bucy filtering problem. Information however evolves exogenously: in their baseline case, these flows are driven by Brownian motions and a Poisson shock. In contrast, our model considers firms who can affect their set of information. As a result, the flow of information becomes an *endogenous* object.

3.3 Stationary Equilibrium

3.3.1 Aggregate Price Consistency

We assume that every firm starts out with the prior λ_0, thus firms are ex-ante homogeneous. However, different realizations of the taste shocks α_k lead to ex-post heterogeneity of a firm's prior belief λ in the cross-section. Furthermore, firm's are ex-post heterogeneous due to different realizations of the idiosyncratic shock in the cross-section. Note there is not only dispersion in the beliefs λ across the firm types but also within firms in each basket. This dispersion in firms' beliefs and their idiosyncratic productivity is captured by the cross-sectional distribution $\varphi_i(\lambda, z)$ for firms of type i. We previously defined the aggregate price index as:

$$\overline{P}_i = \left(\int_{k \in J_i} \alpha_k P_i(k)^{1-\sigma_i} dk \right)^{\frac{1}{1-\sigma_i}}$$

Recall however that the optimal pricing policy $P^*(\lambda, z)$ is independent from j and k. To obtain price consistency in the aggregate, we thus require:

$$\begin{aligned} \overline{P}_i &= \left(\int_{k \in J_i} \alpha_k P_i(k)^{1-\sigma_i} dk \right)^{\frac{1}{1-\sigma_i}} \\ &= \left(\int P^*(\lambda, z)^{1-\sigma_i} d\varphi_i(\lambda, z) \right)^{\frac{1}{1-\sigma_i}} \end{aligned} \tag{6}$$

where we applied the normalization $\int_0^1 \alpha_k dk = 1$ once again.

3.3.2 Labor Market Clearing

The market clearing condition for goods is explicitly incorporated in the firm's problem, thus the only remaining factor market to clear is the labor market. The limit $\xi \to \infty$ for the Frisch elasticity and separable, additive utility in consumption and leisure imply that wages are proportional to aggregate spending.[37] This gives us:

$$W = \omega S$$

[37] See Golosov and Lucas (2007) who use the same specification for consumer preferences. In their setup, this implies that wages are proportional to the stock of money. Thus, it grows at the same rate as inflation. We have a similar proportionality rule as wages become proportional to total spending which grows at the rate of inflation $\tilde{\pi}$.

Nominal total spending S equals $\overline{P}C$ and thus gives us an expression for the *real* wage rate:

$$W/\overline{P} = \omega C$$

Given our linear production technology, labor demand is simply characterized by:

$$H^d = \sum_{i=1}^{2} \int_{k \in J_i} \frac{c_i(k)}{z} dk$$

Plugging in consumer demand (1), total nominal spending and aggregate price consistency (5), we obtain:

$$H^d = S\left[\eta\left(\frac{\int \frac{1}{z}\left[P^*(\lambda,z)\right]^{-\sigma_1} d\varphi_1(\lambda,z)}{\int \left[P^*(\lambda,z)\right]^{1-\sigma_1} d\varphi_1(\lambda,z)}\right) + (1-\eta)\left(\frac{\int \frac{1}{z}\left[P^*(\lambda,z)\right]^{-\sigma_2} d\varphi_2(\lambda,z)}{\int \left[P^*(\lambda,z)\right]^{1-\sigma_2} d\varphi_2(\lambda,z)}\right)\right]$$

Imposing labor market clearing implies $H^s = H^d = H$.

3.3.3 Equilibrium

In the rest of our analysis, we will focus on a *stationary* equilibrium in which any dying firm will be immediately replaced by a new firm. The latter will be assigned to be a type 1 firm with probability λ_0 which will also serve as its prior upon entry. We simplify the analysis by normalizing the measure of firms to 1 and organizing the industry composition as follows:

$$J_1 = [0, \gamma_1] \text{ and } J_2 = (\gamma_1, 1].$$

Our restrictions on entry then imply $\gamma_1 = \lambda_0$ which guarantees a balanced measure of in- and out-flows at the product-level.[38] Whenever nominal total spending grows deterministically at a rate $\tilde{\pi}$, then there is no aggregate uncertainty. Let W be the economy's numéraire and then we can define a stationary equilibrium.

Definition 1 (Stationary equilibrium) *A stationary equilibrium is a tuple* $(W, \overline{P}_1, \overline{P}_2, S)$ *and a pair of invariant distributions* $(\varphi_1(\lambda,z), \varphi_2(\lambda,z))$ *such that real variables are constant. This entails that:*

I. *Consumers maximize utility by consuming varieties* $c_i(k)$*,* $k \in J_i$*,* $i \in \{1,2\}$*.*

II. *Firms engage in optimal pricing strategies and adopt* $P^*(\lambda,z)$ *when adjusting prices.*

III. *Factor markets clear.*

[38]This is equivalent to the balance-of-flows restriction in BM whenever there is *no* endogenous exit.

IV. *Prices are consistently aggregated.*

V. *Firms die at the rate δ and enter the economy as a type 1 firm with probability λ_0.*

In the appendix, we describe the numerical algorithm to solve this framework computationally.

3.4 Characterization of Price Experimentation

To display the price experimentation mechanics as clearly as possible, we will temporarily strip down the model to its most basic version in which menu costs and idiosyncratic cost shocks are discarded and there are only two periods. The monopolist does not know the elasticity of demand is facing but knows that there are two possibilities σ_1 and σ_2.

Two period setup. In the second period, the firm only cares about maximizing myopic profits. Thus:

$$\begin{aligned} V_2(\lambda') &= \max_{p\in\mathcal{P}} \left\{ \left(p - \frac{W}{z}\right) \left(\lambda'\eta \frac{p^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda')(1-\eta)\frac{p^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}}\right) \frac{S}{P} \right\} \\ &\equiv \max_{p\in\mathcal{P}} \mathcal{M}(p;\lambda') \end{aligned}$$

In the first period, the firm must balance between obtaining higher myopic profits and sharpening its posterior beliefs for the next period.

$$\begin{aligned} V_1(\lambda_0) &= \mathcal{M}(p;\lambda_0) + \beta\left[\lambda_0 \int_\varepsilon V_2(b_1(\lambda_0, \log(p), \varepsilon))dF(\varepsilon) + (1-\lambda_0)\int_\varepsilon V_2(b_2(\lambda_0, \log(p), \varepsilon))dF(\varepsilon)\right]\Bigg\} \\ &\equiv \max_{p\in\mathcal{P}} \{\mathcal{M}(p;\lambda_0) + \mathcal{V}(p;\lambda_0)\} \end{aligned}$$

There is no closed form expression for the optimal *myopic* policy function $P^M(\lambda; z)$ under iso-elastic preferences. Hence, we will implicitly define this *myopic* price as:

$$P^M(\lambda) \equiv \arg\max_{P\geq 0} \left\{ \lambda\eta\Pi_1(P) + (1-\eta)(1-\lambda)\Pi_2(P) \right\}$$

Nonetheless, it is straightforward to show that, for a given level of z, $P^M(\lambda)$ is monotonically increasing in λ and can be written as a linear combination of the complete information optimal prices $P_1^*(z)$ and $P_2^*(z)$. This is formalized in the proposition below.

Proposition 1. The myopic policy function $P^M(\lambda)$ is strictly increasing and $\mathcal{C}$ in λ.

Proof. See Appendix A2.2.1. □

The policy $P^M(\lambda)$ would be the optimal pricing function whenever the firm would be unable to affect its posterior beliefs. A firm is said to price *experiment* at the belief λ if it deviates from the myopic price $P^M(\lambda)$.[39] This deviation then exactly reflects the firm's incentive to gain information at the expense of its current period profits.

Convexity of the value of information. As noted before, a firm chooses its price optimally to gain information about its type at the expense of its current profits. As a result, it is obvious that a firm will price experiment if and only if information is valuable. A relatively large literature has established that this is formally captured by a continuation value that is convex in a firm's beliefs.[40] The following lemma establishes this feature.

Lemma 1. The value function $V_2(\lambda)$ is convex and $\mathcal{C}^2$ in λ.

Proof. See Appendix A1.1. □

The result is shown explicitly for the two period setup. However, it is easily generalizable to the infinite period framework. Even though the convexity of $V_2(\lambda)$ is a fairly trivial point to show, we discuss it for two reasons. First, to establish sufficient conditions for experimentation, we follow Mirman et al. (1993). In their proposition 1, the convexitity of $V_2(\cdot)$ is mentioned as one of their two sufficient conditions. Informally, the second condition states that adjustments in prices must be capable of increasing the informativeness of a firm's sales. This will be shown in the proof of proposition 2 and discussed more extensively below.

Incentives to experiment. Recall that a firm does not know its type, but is aware of the fact that its type either has to be $\sigma = \sigma_1$ or $\sigma = \sigma_2$. As a result, a firm sets its price to identify its elasticity of substitution by "separating" these two possible demand curves as much as possible. This implies that the price at which the demand curves cross results in sales that are completely *uninformative*. It is straightforward to deduce that the *expected* demand curves cross if and only if $p = \exp\left(\frac{\Delta\mu}{\Delta\sigma}\right)$. Let this intersecting price be defined as the *confounding* price $\widehat{P}$. If the firm decides to choose its experimentation policy $P^*(\lambda)$ to be equal to $\widehat{P}$, then we expect that there are no benefits of experimentation in this case. As a result, a firm's experimentation policy should coincide with its myopic policy function. This intuition is formalized in proposition 2.

[39] A similar definition can be found in Keller and Rady (1999).

[40] For example, this argument can be found in Aghion et al. (1991).

Proposition 2. Let $\widehat{P} = \frac{\Delta\mu}{\Delta\sigma} \in (P_2^*, P_1^*)$, then there exists a confounding belief $\widehat{\lambda}$ such that:

$$P^*(\widehat{\lambda}) = P^M(\widehat{\lambda}) = \widehat{P}.$$

Furthermore, the confounding belief $\widehat{\lambda}$ is unique up to $\lambda \in \{0, 1\}$ and strictly increasing (decreasing) in $\Delta\mu$ ($\Delta\sigma$).

Proof. See Appendix A.2.2. □

Bounded price experimentation. A firm would set the standard CES mark up over its marginal cost under perfect information, i.e. $P_i^* = \frac{\sigma_i}{\sigma_i - 1}\frac{W}{z}$ for $i = 1, 2$. As a result, it could be conjectured that a firm sets some price $p \in [P_2^*, P_1^*]$ under its experimentation regime. At the same time however, prices that are larger in absolute value are more informative as this induces a larger change in the firm's posterior beliefs.[41] Thus, it is ex-ante unclear whether the firm's pricing space should be bounded. We will deal with this problem by imposing parameter restrictions such that the pricing space under experimentation is bounded. In particular, proposition 3 will provide a sufficient condition such that $\mathcal{P} = [P_2^*, P_1^*]$. The following lemma is used to prove this result.

Lemma 2. The marginal expected change in a firm's posterior belief is bounded in absolute value, i.e.

$$\left| \mathbb{E}_\varepsilon \left[\frac{\partial b_i(\lambda_0, \log(p), \varepsilon)}{\partial \log(p)} \;\middle|\; \varepsilon \in \mathcal{F} \right] \right| \leq \frac{\Delta\sigma}{\sigma_\varepsilon^2} \left(\int_{\varepsilon \in \mathcal{F}} |\log(p)\Delta\sigma - \Delta\mu + \varepsilon| dF(\varepsilon) \right)$$

where the sign of $\log(p)\Delta\sigma - \Delta\mu + \varepsilon$ is constant for all $\varepsilon \in \mathcal{F} \subseteq \mathbb{R}$.

Proof. See Appendix A.1.2. □

Intuitively, the lemma states that the expected change in a firm's posterior beliefs, given some finite price p, can be bounded. Lemma 2 is mainly technical in nature, but will nevertheless be extremely useful for the proof of proposition 3 which is stated below.

[41] Strictly speaking, a firm's demand function can be rewritten such that a firm's pricing experimentation is interpreted as the pure control of variance of its observation process. In other words, larger price changes translate in more informative sales signals through the reduction of the noise caused by the demand shocks. This interpretation of experimentation is also adopted by Moscarini and Smith (2001).

Proposition 3. Let $\frac{\Delta\mu}{\Delta\sigma} \in [P_2^*, P_1^*]$, then:

$$x_{\min} \equiv \log(P_2^*)\Delta\sigma - \Delta\mu - \frac{2\varphi(0)}{1-\Phi\left(-\frac{\log(P_2^*)\Delta\sigma-\Delta\mu}{\sigma_\varepsilon}\right)} < 0,$$

$$x_{\max} \equiv \log(P_1^*)\Delta\sigma - \Delta\mu + \frac{2\varphi(0)}{\Phi\left(-\frac{\log(P_1^*)\Delta\sigma-\Delta\mu}{\sigma_\varepsilon}\right)} > 0.$$

Then, the firm's experimentation policy has an interior solution, i.e. $P^*(\lambda_0) \in (P_2^*, P_1^*)$ for all $\lambda_0 \in (0,1)$ whenever $V_2'(1) < \overline{\mathcal{V}}$ with:

$$\overline{\mathcal{V}} \equiv \frac{\sigma_\zeta^2}{\beta\Delta\sigma} \min\left\{\frac{\eta\lambda_0\Pi_1'(P_2^*)}{-x_{\min}}, \frac{(1-\eta)(1-\lambda_0)\left(-\Pi_2'(P_1^*)\right)}{x_{\max}}\right\}. \tag{7}$$

Proof. See Appendix A.2.3. □

A firm's posterior beliefs are more affected by large prices. Proposition 3 establishes however that a firm's experimentation policy is bounded from below and above by the perfect information extremes. The stated condition in the proposition above basically implies that the *net* marginal benefits are strictly positive at $p = P_2^*$ and vice versa for $p = P_1^*$ and a natural way to satisfy it is for either $\Pi_1'(P_2^*)$ or $-\Pi_2'(P_1^*)$ to be sufficiently large. This occurs whenever the static profit functions $\Pi_i(p)$ for $i = 1, 2$ possess sufficient curvature. The result in proposition 3 is then intuitive: whenever the curvature of a firm's profit function is sufficiently high, deviations from the myopic price become more costly. This is equivalent to saying that the opportunity costs from experimentation become higher. As a result, a firm will never experiment in a way that is "too extreme".

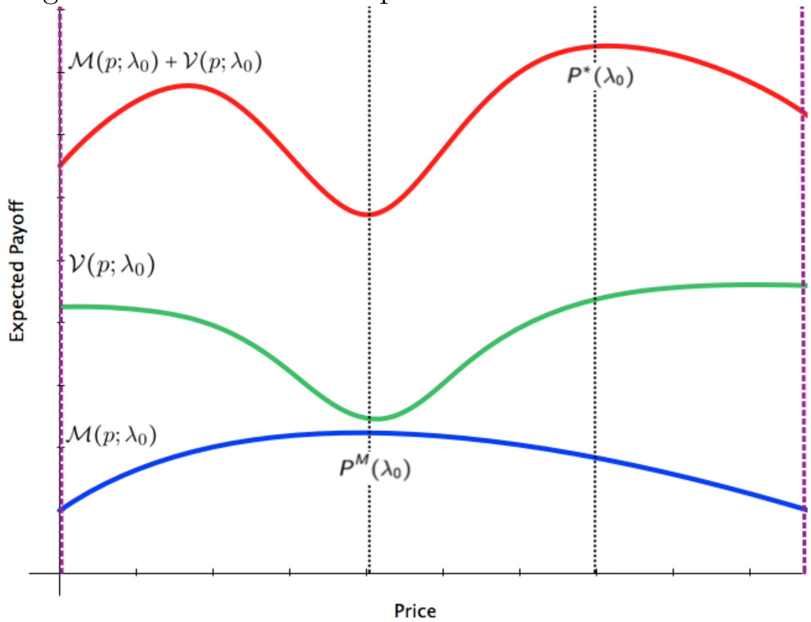
Numerical example. Figure 5 plots the firm's static profits $\mathcal{M}(p;\lambda_0)$, the continuation value $\mathcal{V}(p;\lambda_0)$, and the total payoff $\mathcal{M}(p;\lambda_0) + \mathcal{V}(p;\lambda_0)$ as a function of price.[42] The dotted lines at the extremes of the figure depict the optimal prices P_2^* and P_1^*. Given that $\mathcal{M}(p;\lambda_0)$ is a weighted sum of strictly concave functions in p, it is strictly concave and is maximized at $P^M(\lambda_0)$. The concavity of $\mathcal{M}(p;\lambda_0)$ illustrate the costs of experimentation as prices far from $P^M(\lambda_0)$ represent profit losses in the first period.

The gains from experimentation are represented by the continuation value $\mathcal{V}(p;\lambda_0)$ which is convex and its minimum is at the confounding price; the price at which a firms sales are rendered as completely uninformative.[43] In this case, small deviations from the confounding

[42]For simplicity, in this example $\lambda_0 = \widehat{\lambda}$.

[43]The convexity of $\mathcal{V}(p;\lambda_0)$ reflects back to the second sufficient condition in Mirman et al. (1993) which states that adjustments in prices must be capable of generating additional information by making a firm's

Figure 5: Numerical Example of the Two Period Model



Note: Static profits $\mathcal{M}(p;\lambda_0)$, continuation value $\mathcal{V}(p;\lambda_0)$ and total payoff $\mathcal{M}(p;\lambda_0)+\mathcal{V}(p;\lambda_0)$ at $\lambda_0=\widehat{\lambda}$. The dotted purple lines represent the optimal prices P_2^* and P_1^*. $P^M(\lambda_0)$ represents the myopic policy and $P^*(\lambda_0)$ the policy under experimentation.

price lead to large gains. This is, the benefits from experimentation are strongly related to the convexity of $\mathcal{V}(p;\lambda_0)$. For example, prior beliefs closer to 0 and 1 lead to less convex continuation values. This is intuitive as the marginal benefit of information decreases for firms that are more certain about their type and, as a result, the benefits from experimentation are reduced. The convexity of $\mathcal{V}(p;\lambda_0)$ is also affected by the signal-to-noise ratio. For extremely large values of σ_ϵ, for instance, the optimal policy converges to the myopic policy. The convexity of $\mathcal{V}(p;\lambda_0)$ will be the most important determinant of the different experimentation regimes that we discuss below.

Lastly, in this example, the sum of a strictly concave and convex function results in a function that is double-peaked and whose local maximum is at $P^*(\lambda_0)$. The figure shows that the global maximum lies in the interior of $\mathcal{P}=[P_2^*,P_1^*]$ which is what is shown in proposition 3. The degree of experimentation is captured by the difference between $P^*(\lambda_0)$ and $P^M(\lambda_0)$.

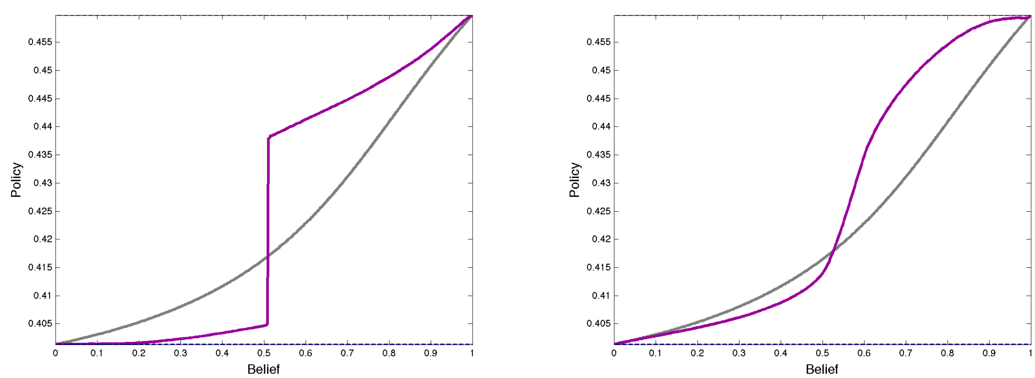
Experimentation regimes. The gains from experimentation and strongly related to the convexity of $\mathcal{V}(p;\lambda_0)$. This, in turn, is determined by the prior belief, the signal-to-noise ratio, and the discount factor. The prior belief determines how certain the firm is about its type. The gains from experimentation are reduced as the firm has more certainty about the elasticity of demand it is facing. The signal-to-noise ratio summarizes the sensitivity of the

sales signal more informative.

posterior beliefs to price deviations relative to the crossing point. Firms that face extremely large levels of noise will not find profitable to experiment. And, lastly, the discount factor indicates who much the firm values the future and, as a result, who much it values information. The convexity of $\mathcal{V}(p;\lambda_0)$ determines the shape of the total payoff function. In our previous numerical example, the total payoff function was double-peaked because $\mathcal{V}(p;\lambda_0)$ had enough convexity but this might not always be the case.

The shape of the total payoff determines the experimentation regime.[44] In our setup, there are two qualitatively different regimes determined by the shape of $\mathcal{V}(p;\lambda_0)$: extreme and moderate experimentation. Under the extreme experimentation regime, the total payoff function is double-peaked. As a result, the firm never chooses to price at the confounding price and $P^*(\lambda)$ displays a discontinuity at $\lambda = \widehat{\lambda}$. Intuitively, since the value of information is minimized at the confounding belief $\widehat{\lambda}$, the firm has most incentives to change its price at this specific belief and deviates in a discontinuous fashion. Under moderate experimentation, the total payoff function is single-peaked and the policy function $P^*(\cdot)$ is continuous between P_2^* and P_1^*.[45]

Figure 6: Experimentation Regimes



(a) Extreme Experimentation. (b) Moderate Experimentation.

Note: Panel (a) shows the extreme experimentation regime and panel (b) the moderate. The gray line depicts the myopic policy and the purple lines the policy under experimentation. The dotted lines at the top and bottom of the panels indicate the optimal prices P_1^* and P_2^*. The price at which the myopic policy and the experimentation policy cross is the confounding price.

Figure 6 depicts the two experimentation regimes. The thin gray line shows the myopic policy function $P^M(\lambda)$ which is monotonically increasing in λ whereas the purple line is the policy function $P^*(\lambda)$ under experimentation. Consistent with proposition 3, we observe that $P^M(\lambda)$ and $P^*(\lambda)$ are bounded from below and above by P_2^* and P_1^* respectively. Under ex-

[44]The terminology is borrowed from Keller and Rady (1999) who show that different experimentation regimes could arise in a problem of a seller choosing quantities subject to a randomly changing state

[45]This will be formalized in a future version of this draft.

treme experimentation, the policy function shows a discontinuity at the confounding belief. The firm experiments mostly near $\widehat{\lambda}$ as it tries to keep the informativeness of its observed sales as high as possible. Obviously, it can only do this to a limited extend as otherwise the firm would lose too many static profits. Under the moderate experimentation regime, the myopic and the experimentation policies coincide at the confounding price $\widehat{P}$ as predicted by proposition 2. Once the firm updates its posterior closer to the boundaries (i.e. $\lambda = 0$ or $\lambda = 1$), the incentives for experimentation decline again as the firm's information set converges to the complete information case. In this case, the myopic and experimentation policies coincide at $\lambda \in \{0, 1\}$. Hence, the firm would never pay the opportunity costs (i.e. give up static profits) through experimentation whenever beliefs reach either 0 or 1.[46]

[46] BM argue that price experimentation imparts an *upward* bias in prices though. According to them, a high price has the benefit of reducing physical sales and hence production costs. This principle goes beyond iso-elastic preferences. However, we conjecture that this is not always the case. More importantly, this intuition is extremely sensitive to the underlying calibration. This is illustrated in figure **??** and **??** below for example.

4 Quantifying Monetary Non-neutrality with Price Experimentation

4.1 Calibration and Results

The IRI Symphony data is measured at the weekly level, so we set the model period to be one week. As a result, the discount factor is set at $\beta = 0.96^{1/52}$ which reflects an interest rate of around 3.8% and incorporates the exogenous exit rate of 0.4% which is directly taken from the IRI Symphony data at the UPC-store level.

The mean yearly growth rates of nominal and real GDP equal $g_n = 0.04$ and $g_r = 0.02$ respectively. Since there is no long-run real growth in the model economy, we set $\tilde{\pi} = (g_n - g_r)^{1/52} = 0.00038$ as the weekly rate of inflation. Furthermore, standard deviation of the taste shock σ_ε equals 0.4 which matches the standard deviation of sold quantities conditional on no price change in the IRI Symphony.[47] Lastly, the disutility of labor is chosen so that aggregate employment is approximately $\frac{1}{3}$ as we normalized the amount of time available to the consumer to unity.

The remaining parameters are calibrated to match various micro data moments. There are seven remaining parameters: two elasticities of substitution (σ_1 and σ_2), prior belief upon entry λ_0, basket division of income η, fixed menu cost ψ and the persistence and standard deviation of idiosyncratic productivity ρ and σ_ς. These parameters are calibrated jointly and are selected to hit eight moments from the data: the average frequency of adjustment, average size of increases, average size of decreases, the fraction of price changes that are increases, the frequency of adjustment on the second and tenth week since entry and the absolute size of price changes during the second and tenth week since entry.

Table I: Internally Calibrated Values of the Model's Parameters

Description	Parameter	Value
Elasticity of Substitution 1	σ_1	4.7
Elasticity of Substitution 2	σ_2	13.4
Prior Belief at Entry	λ_0	0.75
Basket Division of Income	η	0.36
Fixed Cost	ψ	0.02
Productivity Persistence	ρ	0.56
Productivity Standard Deviation	σ_ς	0.05

As is standard in the class menu cost models, the fixed cost of adjustment ψ will partially governs the average frequency and size of adjustments. The extent to which price experimen-

[47]Our findings are consistent with the value of 42 percent which is reported in Eichenbaum et al. (2008).

tation is more present at the beginning of a product's life-cycle is determined by the amount of information a firm has at entry summarized by the signal-to-noise ratio $(\sigma_2 - \sigma_1)/\sigma_\varepsilon$ and the prior belief upon entry λ_0 which represents the fraction of firms facing elasticity of substitution σ_1. We assume this parameter to be equal across all entering firms. As the incentives to experiment decrease, price changes are mainly driven driven by idiosyncratic cost shocks.[48] Thus, the parameters ρ and σ_ζ will have a relatively large impact on the pricing moments at the later stage of the product life-cycle.

Table II: Moments of Price Change Distribution

Moment	Data	Model with Learning
Frequency	0.05	0.05
Fraction Up	0.66	0.57
Size Up	0.09	0.09
Size Down	0.07	0.11
Frequency Week 2	0.09	0.09
Frequency Week 10	0.06	0.05
Absolute Size Week 2	0.17	0.17
Absolute Size Week 10	0.11	0.10
Moments Not Targeted		
Std. of Price Changes	0.11	0.10
75th Pct Size Price Changes	0.10	0.11
90th Pct Size Price Changes	0.18	0.13

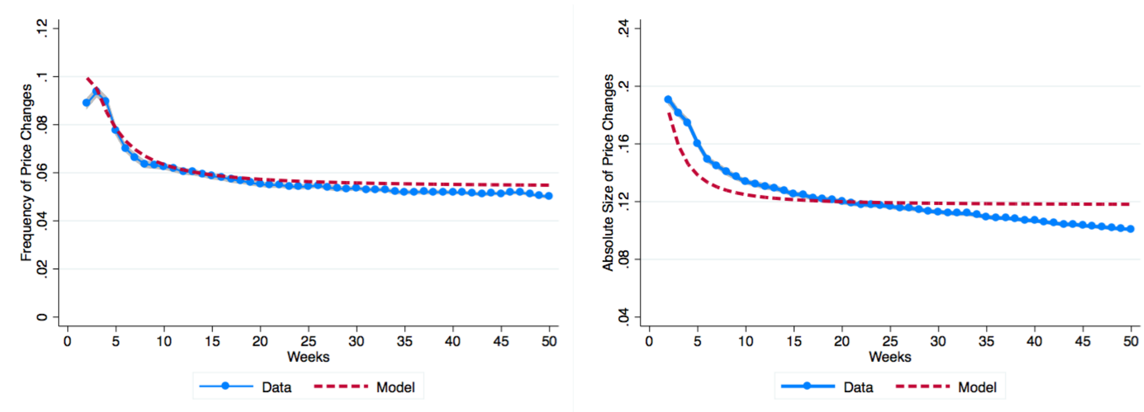
Table I shows the model's best parameters in terms of fitting moments and table II displays the resulting moments from the framework compared to the data. The productivity parameters are in line with previous estimates in the menu cost literature. The model matches the frequency of adjustment and fraction of increasing price changes quite precisely. The specification for the menu cost ψ implies that the total adjustment costs in the economy represent approximately 0.7% of steady-state weekly revenues. The cost conditional on adjustment is around 1.4% which is in line with the estimates in Zbaracki et al. (2004). The value of σ_1 is in the range of values typically used in the menu cost literature. The model requires somewhat larger σ_2 to induce enough experimentation. Nonetheless, σ_2 is well within the estimates of Broda and Weinstein (2010) who compute elasticities of substitution for a variety of products using data similar to ours.

The model also does a good job of replicating the life-cycle patterns of the frequency of price adjustment and the absolute size of adjustment qualitatively. Figure 7 shows that in our simulations entering products are more likely to adjust prices and they do so by larger

[48]This is seen most clearly from the Barro-Dixit-Stiglitz pricing formula in which the range of inaction is characterized by some interval $[-\overline{p}, \overline{p}]$ where $\overline{p} = \left(6\frac{\sigma_\zeta^2 \psi}{B}\right)^{1/4}$.

amounts. This is driven by 1) the size of the signal-to-noise ratio and 2) the overall *level* of the elasticities of substitution since firms with lower market power (i.e. high elasticities of substitution) have higher incentives to get their prices "right" as the opportunity costs of experimentation (i.e. sacrificing static profits) are higher.

Figure 7: Model vs Data



(a) Frequency of Price Changes (b) Absolute Size of Adjustments

Note: The figure shows the results of the model and compares them with the data. We simulate a panel of 1000 firms over 1000 periods and compute both the predicted frequency of price adjustment and the absolute size of price changes over the life-cycle of a product. The results of the *frequency* of price changes are shown in panel (a) and those of the *absolute size* of price changes are shown in panel (b).

The incentives to experiment also affect the size distribution of price changes by generating large price changes *endogenously*. Our calibration matches the standard deviation of price changes and price changes in the 75th percentile in absolute value without explicitly targeting them. However, the model under-predicts the prevalence of price changes in the 90th percentile of the size-distribution. Furthermore, the resulting hazard of price changes in our economy is downward sloping with a hump at short durations. This is not obvious at first glance and it is the result of two opposing forces. As a result of price experimentation, the probability of consecutive price changes is larger at entry. As demand uncertainty declines also does the probability of price adjustment, generating a decreasing hazard rate. This force is proportional to the number of firms in the economy conducting price experimentation. On the other hand, the presence of a menu costs make firms less likely to adjust after they reset their prices, generating an upward sloping hazard rate.

4.2 Implications

Price experimentation has important implications for the transmission of nominal shocks to the real economy. We perform a counterfactual experiment in which log nominal output increases permanently by a size that is comparable to a one week doubling of the nominal output growth rate. We observe that about 70% of the nominal shock goes into output *on impact*. A baseline menu cost model with full information in the spirit of Golosov and Lucas (2007) features a value of around 60%, i.e. only around 60% of the nominal shock goes into real output on impact in a model without experimentation.[49]

To provide intuition for this result, we follow Caballero and Engel (2007) and decompose the price response on impact in the class of Ss models into intensive and extensive margins. Let $x_t(\lambda) \equiv \log\left(\frac{p_t}{p_t^*(\lambda)}\right)$ be the difference between a firm's current price and its *desired* price, i.e. the price it will choose as a function of its beliefs conditional on adjusting. $x_t(\lambda)$ is also known as the *price gap*. Let the economy-wide distribution of price gaps be given by $f(x, \lambda)$ and assume that firms have an adjustment probability that is increasing in their price gap $\Lambda(x, \lambda)$.[50] This implies that inflation will be given by:

$$\pi = \int x(\lambda)\Lambda(x, \lambda) f(x, \lambda) dx d\lambda$$

If there is some unexpected, positive shock $\triangle S > 0$ to firms' desired prices, inflation will be characterized by:

$$\pi(\triangle S) = \int (x(\lambda) + \triangle S)\Lambda(x + \triangle S, \lambda) f(x, \lambda) dx d\lambda$$

Taking a first-order Taylor approximation of $\pi(\triangle S)$ around $\triangle S = 0$, rearranging and taking the limit as $\triangle S \to 0$ gives that the price response on impact equals:

$$\lim_{\triangle S \to 0} \frac{\triangle \pi}{\triangle S} = \underbrace{\int \Lambda(x, \lambda) f(x, \lambda) dx d\lambda}_{= \text{intensive}} + \underbrace{\int x(\lambda)\Lambda_x(x, \lambda) f(x, \lambda) dx d\lambda}_{= \text{extensive}} \tag{8}$$

This price response on impact is the sum of two components. First, the intensive margin gives the inflation contribution of firms' products whose prices would have adjusted without the aggregate shock. These firms adjust to the aggregate shock by changing the *size* of their adjustment. Equation 8 shows that this margin equals the frequency of adjustment. Second, the extensive margin captures the strength of the selection effect and gives the additional

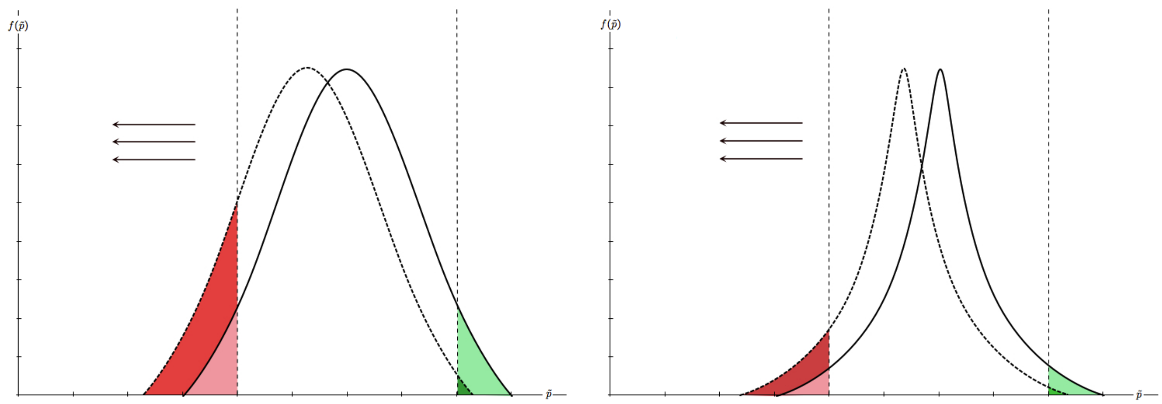
[49]The full information model in this case also features two types of firms (those facing σ_1 and those facing σ_2) but they do not have uncertainty about their type. This model is calibrated to match the same moments as the model with learning and features the same fraction of firms of each type.

[50]To simplify the math, we assume here that a positive shock $\triangle S$ does not affect the firm's belief. However, our results reported below *do* take this effect in consideration as we calculate the extensive and intensive margins numerically.

inflation contribution of firms whose decision to adjust is either triggered or canceled by the aggregate shock. The extensive margin becomes naturally more relevant as the number of firms on or near the margin of adjustment increases. This is reflected by a relatively large value for $\Lambda_x(x, \lambda)$. Furthermore, it is more important whenever the difference between adjusting and not adjusting is large which is equivalent to a large, absolute price gap $|x(\lambda)|$.

In the Golosov-Lucas model, the ratio of the extensive to the intensive margin is between 4 and 5.[51]. On the other hand, since the experimentation motive changes the distribution of desired price changes, there is a *lower* mass of firms at the original bounds of inaction. Under price experimentation, the fraction of firms at the original bounds is smaller. As a result, the extensive margin becomes less important. This is reflected in our numerical simulations as the ratio of extensive to intensive margin is approximately 3.

Figure 8: Distribution of Desired Price Changes



(a) Golosov-Lucas (b) Experimentation

Note: The figure plots steady state desired price change distributions of the Golosov-Lucas model and a model with experimentation. The shaded areas show the price change distribution in response to an aggregate shock.

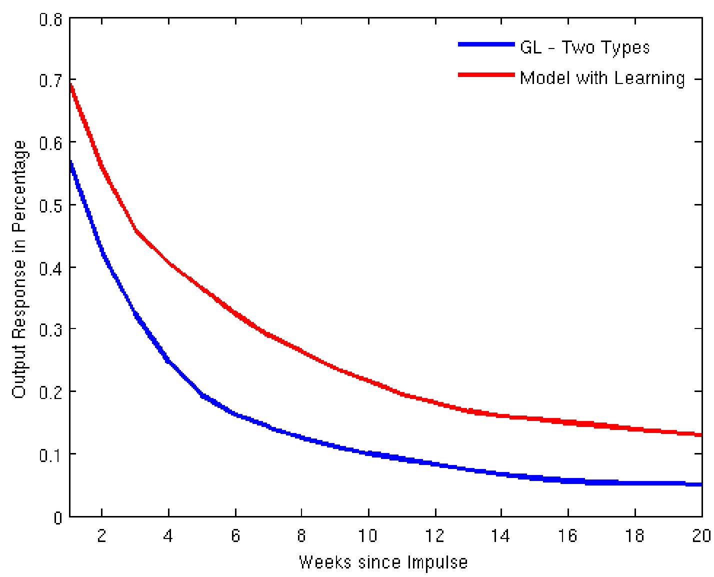
The intuition of the impact of price experimentation is captured in figure 8. The aggregate shock decreases each firm's desired price change and moves the distribution to the left relative to the inaction band. As shown in the figure, the fraction of firms adjusting prices in response to the shock in our setup is a lot smaller than in the Golosov-Lucas model. Therefore, despite the fact that the contribution of the intensive margin is equal in both models, the extensive margin decreases under a setup with price experimentation. This result is not obvious at first glance and is the result of two opposing forces. On the one hand, price experimentation leads to larger desired price gaps $|x(\lambda)|$ which increases the extensive margin. On the other hand, the desire to experiment with prices pushes firms away from the margin of adjustment (which is equivalent to a lower $\Lambda_x(x, \lambda)$) and hence decreases the extensive margin. In our

[51]In the Calvo model, the extensive margin is zero as there is no selection effect at all in this case.

calibration, we find that the latter effect is stronger.

In addition to these differences on impact, the half-life of the real response almost doubles with respect to that of the Golosov-Lucas model. This is because, by introducing the notion of a product's life-cycle, we introduced cross-sectional heterogeneity in the frequency of price adjustment across firms of different ages. Experimenting firms have vastly higher frequencies of price change. Thus, these firms will adjust their price most likely several times before firms with sharper beliefs have adjusted their price even once. However, as pointed out by Nakamura and Steinsson (2008b), all price changes after the first one for the experimenting firms do not affect output on average because these firms have already adjusted to the shock. Given that the model is calibrated to match the *average* frequency of price changes, the fact that firms who are certain about their type have on average lower frequency of adjustment significantly delays the adjustment of the aggregate price level after a nominal shock.[52]

Figure 9: Real Output Response to Nominal Shock



Note: The figure shows the response of log real output to a 0.00038 increase in the nominal output growth rate. The output response is shown in the graph as a percent of the nominal shock. The red line depicts the output response in Golosov and Lucas (2007) with two different types of firms (i.e. σ_1 and $sigma_2$) and the blue line the response in a price-setting model with price experimentation. Both models are calibrated to match the same moments and feature the same fraction of firms of each type.

As a result, the cumulative effects on real output are larger under demand uncertainty as

[52]This can be illustrated in a simple Calvo model. In that framework, the degree of monetary-non-neutrality is convex in the frequency of price changes. If, for example, the overall frequency of price adjustment in the economy is a convex function of the frequency of price changes of firms experimenting and those certain about the elasticity they face, heterogeneity in the cross-sectional distribution of firms will amplify monetary non-neutrality.

shown in Figure 9. The cumulative increase in log real output to log nominal output shock of size 0.000381 is approximately 0.002 so that real output is increased by 516% relative to the size of the nominal shock. This implies that the cumulative output impulse response increases approximately 40% larger than in the model with full information.

5 Conclusion

In this paper, we documented two novel facts on pricing moments over the life-cycle of U.S. products and provided a structural interpretation for it through a price experimentation mechanism. We find that entering products are not only subject to significantly higher frequencies of price changes (two times higher) but also experience substantially larger adjustments (50 percent larger) relative to the average.

To account for these novel set of facts, we use a menu cost model with price experimentation. We develop the conditions for experimentation and describe the different regimes that could arise in our context. This approach has two additional benefits. First, the framework is able to replicate the stylized facts over the life-cycle of products and it generates large price changes as consistent in the data *without* the use of fat-tailed shocks. To the best of our knowledge, our framework is the first to explain all of these facts *simultaneously*.

Our calibrated framework, which can be interpreted as a hybrid between standard menu cost models à la Golosov and Lucas (2007) and price experimentation in the tradition of Mirman et al. (1993) and Bachmann and Moscarini (2012), has important implications for monetary non-neutrality. Relative to the Golosov-Lucas benchmark, we show that the direct impact of a nominal money shock is higher. Furthermore, the real effects of nominal shocks become more persistent as the half-life response is increased by roughly one-fifth and the cumulative output effect are 40% larger.

While we made significant progress on the quantitative implications of price experimentation in a menu cost model, there are still many areas worthwhile to pursue. In particular, there is progress to be made on *analytically* characterizing the conditions that led to the different experimentation regimes. Furthermore, we believe that a sharper characterization on the decomposition of the price response on impact between intensive and extensive margins is possible. While we have characterized this decomposition quantitatively, theoretical results could greatly increase our understanding of the impact of uncertainty on the effectiveness of monetary policy.

Lastly, it would be interesting to explore the robustness of our results to the *type* of experimentation. We have chosen a fairly parsimonious approach to price experimentation. However, the implications of uncertainty and experimentation through optimal filtering and control problems are also worthwhile to analyze.

References

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien, "Optimal learning by experimentation," *The review of economic studies*, 1991, *58* (4), 621–654.

Aguiar, Mark and Erik Hurst, "Deconstructing lifecycle expenditure," Technical Report, National Bureau of Economic Research 2008.

Alvarez, Fernando and Francesco Lippi, "Price setting with menu cost for multiproduct firms," *Econometrica*, 2014, *82* (1), 89–135.

Alvarez, Fernando E, Francesco Lippi, and Luigi Paciello, "Optimal Price Setting With Observation and Menu Costs*," *The Quarterly journal of economics*, 2011, *126* (4), 1909–1960.

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi, "Small and large price changes and the propagation of monetary shocks," Technical Report, National Bureau of Economic Research 2014.

Bachmann, Rüdiger and Giuseppe Moscarini, "Business Cycles and Endogenous Uncertainty," 2012.

Baley, Isaac and Julio A Blanco, "Menu costs, uncertainty cycles, and the propagation of nominal shocks," Technical Report 2014.

Barro, Robert J, "A theory of monopolistic price adjustment," *The Review of Economic Studies*, 1972, pp. 17–26.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott, "Firms in international trade," Technical Report, National Bureau of Economic Research 2007.

_ , Stephen J Redding, and Peter K Schott, "Multiple-product firms and product switching," *American Economic Review*, 2010, *100* (1), 70–97.

Broda, Christian and David E Weinstein, "Product Creation and Destruction: Evidence and Price Implications," *The American Economic Review*, 2010, pp. 691–723.

Bronnenberg, Bart J, Michael W Kruger, and Carl F Mela, "Database paper-The IRI marketing data set," *Marketing Science*, 2008, *27* (4), 745–748.

Caballero, Ricardo J and Eduardo MRA Engel, "Price stickiness in Ss models: New interpretations of old results," *Journal of monetary economics*, 2007, *54*, 100–121.

Calvo, Guillermo A, “Staggered prices in a utility-maximizing framework,” *Journal of monetary Economics*, 1983, *12* (3), 383–398.

Campbell, Jeffrey R and Benjamin Eden, “Rigid prices: evidence from US scanner data,” Technical Report, Federal Reserve Bank of Chicago 2005.

Caplin, Andrew and Daniel Spulber, “Menu costs and the neutrality of money,” 1987.

Chevalier, Judith A and Anil K Kashyap, “Best Prices: Price Discrimination and Consumer Substitution,” Technical Report, National Bureau of Economic Research 2014.

__ , __ , and Peter E Rossi, “Why Don’t Prices Rise During Periods of Peak Demand? Evidence from Scanner Data,” *American Economic Review*, 2003, *93* (1), 15–37.

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*, 2015.

Deaton, Angus and Christina Paxson, “Intertemporal choice and inequality,” in “Journal of political economy” Citeseer 1994.

Dixit, Avinash, “Analytical approximations in models of hysteresis,” *The Review of Economic Studies*, 1991, *58* (1), 141–151.

Easley, David and Nicholas M Kiefer, “Controlling a stochastic process with unknown parameters,” *Econometrica: Journal of the Econometric Society*, 1988, pp. 1045–1064.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo, “Reference prices and nominal rigidities,” Technical Report, National Bureau of Economic Research 2008.

Gagnon, Etienne and David Lopez-Salido, “Small price responses to large demand shocks,” 2014.

Gaur, Vishal and Marshall L Fisher, “In-store experiments to determine the impact of price on sales,” *Production and Operations Management*, 2005, *14* (4), 377–387.

Gertler, Mark and John Leahy, “A Phillips Curve with an Ss Foundation,” *Journal of Political Economy*, 2008, *116* (3).

Golosov, Mikhail and Robert E Lucas, “Menu Costs and Phillips Curves,” *Journal of Political Economy*, 2007, *115* (2).

Gourinchas, Pierre-Olivier and Jonathan A Parker, “Consumption over the life cycle,” *Econometrica*, 2002, *70* (1), 47–89.

Karadi, Peter and Adam Reiff, “Large shocks in menu cost models,” 2012.

Kehoe, Patrick J and Virgiliu Midrigan, “Sales and the Real Effects of Monetary Policy,” 2007.

Keller, Godfrey and Sven Rady, “Optimal experimentation in a changing environment,” *The review of economic studies*, 1999, *66* (3), 475–507.

Kiefer, Nicholas M and Yaw Nyarko, “Optimal control of an unknown linear process with learning,” *International Economic Review*, 1989, *30* (3), 571–86.

Kryvtsov, Oleksiy and Nicolas Vincent, “On the Importance of Sales for Aggregate Price Flexibility,” Technical Report, Bank of Canada Working Paper 2014.

Malin, Benjamin A, Eric Anderson, Emi Nakamura, Duncan Simester, and Jon Steinsson, “Informational Rigidities and the Stickiness of Temporary Sales,” Technical Report, Federal Reserve Bank of Minneapolis 2015.

Melitz, Marc J, “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 2003, *71* (6), 1695–1725.

Midrigan, Virgiliu, “Menu costs, multiproduct firms, and aggregate fluctuations,” *Econometrica*, 2011, *79* (4), 1139–1180.

Mirman, Leonard J, Larry Samuelson, and Amparo Urbano, “Monopoly experimentation,” *International Economic Review*, 1993, *34* (3), 549–563.

Moscarini, G. and L. Smith, “The Optimal Level of Experimentation,” *Econometrica*, 2001, *69* (6), 1629–1644.

Nakamura, Emi and Jón Steinsson, “Five facts about prices: A reevaluation of menu cost models,” *The Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

_ and Jon Steinsson, “Monetary non-neutrality in a multi-sector menu cost model,” Technical Report, National Bureau of Economic Research 2008.

Nardi, Mariacristina De, Eric French, and John Bailey Jones, “Why do the elderly save? the role of medical expenses,” Technical Report, National Bureau of Economic Research 2009.

Prescott, Edward C, “The multi-period control problem under uncertainty,” *Econometrica: Journal of the Econometric Society*, 1972, pp. 1043–1058.

Rothschild, Michael, “A two-armed bandit theory of market pricing,” *Journal of Economic Theory*, 1974, *9* (2), 185–202.

Stroebel, Johannes and Joseph Vavra, “House Prices, Local Demand, and Retail Prices,” *Available at SSRN 2500457*, 2014.

Vavra, Joseph, “Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation,” *The Quarterly Journal of Economics*, 2014, *129* (1), 215–258.

Willems, Tim, “Acitvely Learning by Pricing: A Model of an Experimenting Seller,” *Available at SSRN 2381430*, 2013.

Zbaracki, Mark J, Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen, “Managerial and customer costs of price adjustment: direct evidence from industrial markets,” *Review of Economics and Statistics*, 2004, *86* (2), 514–533.

A Appendix

A.1 Proof of Lemmas

A.1.1 Proof of Lemma 1

Proof. Recall that the value function $V_2(\lambda)$ is given by:

$$V_2(\lambda) = \max_{p \in \mathcal{P}} \left\{ \left(p - \frac{W}{z}\right) \left(\lambda\eta \frac{p^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda)(1-\eta)\frac{p^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}} \right) \frac{S}{P} \right\}$$

The function $f(\lambda, p) \equiv \left(p - \frac{W}{z}\right) \left(\lambda\eta \frac{p^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda)(1-\eta)\frac{p^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}} \right) \frac{S}{P}$ is continuous in $(\lambda, p) \in [0,1] \times \mathcal{P}$. The set $\mathcal{P} = [P_2^*, P_1^*]$ is furthermore compact. Then, the *Theorem of the Maximum* implies that $V_2(\cdot)$ is continuous on $[0,1]$.

Convexity in λ follows almost directly. Fix an arbitrary $\alpha \in [0,1]$ and $\lambda, \lambda' \in [0,1]$. Let the convex combination $\tilde{\lambda}$ be defined as $\alpha\lambda + (1-\alpha)\lambda'$ and define the myopic policy function:

$$P^M(\lambda) = \arg\max_{p \in \mathcal{P}} \left\{ \left(p - \frac{W}{z}\right) \left(\lambda\eta \frac{p^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda)(1-\eta)\frac{p^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}} \right) \frac{S}{P} \right\}$$

Then, we get:

$$\begin{aligned}
V_2(\tilde{\lambda}) = \alpha &\left(P^M(\tilde{\lambda}) - \frac{W}{z}\right) \left(\lambda\eta \frac{P^M(\tilde{\lambda})^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda)(1-\eta)\frac{P^M(\tilde{\lambda})^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}} \right) \frac{S}{P} \\
&+ (1-\alpha) \left(P^M(\tilde{\lambda}) - \frac{W}{z}\right) \left(\lambda'\eta \frac{P^M(\tilde{\lambda})^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda')(1-\eta)\frac{P^M(\tilde{\lambda})^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}} \right) \frac{S}{P} \\
&\leq \alpha V_2(\lambda) + (1-\alpha) V_2(\lambda')
\end{aligned}$$

Therefore, we showed $V_2(\alpha\lambda + (1-\alpha)\lambda') \leq \alpha V_2(\lambda) + (1-\alpha)V_2(\lambda')$ which is equivalent to $V_2(\cdot)$ being convex. □

A.1.2 Proof of Lemma 2

Proof. Let $x \equiv \log(P)\Delta\sigma - \Delta\mu$ and ε is contained in some set $\mathcal{F} \subseteq \mathbb{R}$. By construction of the ex-post belief function $b_i(\lambda, \log(P), \varepsilon)$, we obtain:

$$\left| \mathbb{E}_\varepsilon \left[\frac{\partial b_i(\lambda_0, \log(P), \varepsilon)}{\partial \log(P)} \;\middle|\; \varepsilon \in \mathcal{F} \right] \right| = \left| \int_{\varepsilon \in \mathcal{F}} \frac{\exp\left(\frac{(\varepsilon+x)^2+\varepsilon^2}{2\sigma_\varepsilon^2}\right) \Delta\sigma(x+\varepsilon)(1-\lambda_0)\lambda_0}{\left(\exp\left(\frac{\varepsilon^2}{2\sigma_\varepsilon^2}\right)(1-\lambda_0)\sigma_\varepsilon + \exp\left(\frac{(x+\varepsilon)^2}{2\sigma_\varepsilon^2}\right)\lambda_0\sigma_\varepsilon\right)^2} dF(\varepsilon) \right|$$

$$= \frac{\Delta\sigma}{\sigma_\varepsilon^2}\left| \int_{\varepsilon\in\mathcal{F}} \left(\frac{\exp\left(\frac{(\varepsilon+x)^2}{2\sigma_\varepsilon^2}\right)\lambda_0}{\exp\left(\frac{\varepsilon^2}{2\sigma_\varepsilon^2}\right)(1-\lambda_0)+\exp\left(\frac{(\varepsilon+x)^2}{2\sigma_\varepsilon^2}\right)\lambda_0} \right) \times \right.$$
$$\left. \left(\frac{\exp\left(\frac{\varepsilon^2}{2\sigma_\varepsilon^2}\right)(1-\lambda_0)}{\exp\left(\frac{\varepsilon^2}{2\sigma_\varepsilon^2}\right)(1-\lambda_0)+\exp\left(\frac{(\varepsilon+x)^2}{2\sigma_\varepsilon^2}\right)\lambda_0} \right)(x+\varepsilon)dF(\varepsilon) \right|$$
$$\leq \frac{\Delta\sigma}{\sigma_\varepsilon^2}\left(\int_{\varepsilon\in\mathcal{F}} |x+\varepsilon| dF(\varepsilon) \right)$$

where the last inequality follows as the bracketed terms in the second equality are bounded in $[0,1]$ and the sign of $x+\varepsilon$ remains constant on the set $\mathcal{F}$ by assumption. This is exactly what we wanted to show. □

A.2 Proof of Propositions

A.2.1 Proof of Proposition 1

Proof. The *Theorem of the Maximum* implies that $P^M(\lambda)$ is a non-empty, compact-valued and upper hemi-continuous correspondence. However, the objective function is a weighted average of strictly concave functions, thus it is strictly concave itself. As a result, $P^M(\lambda)$ must be single-valued. This implies that $P^M(\lambda)$ is not only upper hemi-continuous but continuous.

Appendix A.1 of Bachmann and Moscarini (2012) implies that $\frac{dP^M(\lambda)}{d\lambda} > 0$ if and only if $P^M(\lambda) > P^M(0) = P_2^*$ for $\lambda > 0$. By construction, this holds for $\lambda = 1$ as $P_1^* > P_2^*$ as $\sigma_2 > \sigma_1$. Thus, the inequality must hold as well for large enough λ through continuity of $P^M(\cdot)$.

Suppose by way of contradiction that for some $\lambda' > 0$, we have that $P^M(\lambda') = P^M(0)$ instead. Then for some small $\Delta > 0$, we must either have $P^M(\lambda'-\Delta) > P^M(0)$, $P^M(\lambda'-\Delta) = P^M(0)$ or $P^M(\lambda'-\Delta) < P^M(0)$. The first case implies that $\frac{dP^M(\lambda)}{d\lambda} < 0$ which contradicts the equivalence from Bachmann and Moscarini (2012). The second case states that $P^M(\lambda'-\Delta) = P^M(0)$ over an open interval of small strictly positive values of Δ. However, this cannot be true as the expected profit function is strictly concave. Whenever $P^M(\lambda'-\Delta) < P^M(0)$, then we must have $\frac{dP^M(\lambda)}{d\lambda}\big|_{\lambda=\ell} < 0$ for all $\ell \in (0,\lambda')$. This implies however that for all $\ell \in (0,\lambda')$ we have $P^M(\ell) < P^M(0)$ but we assumed that $\lim_{\lambda\downarrow 0} P^M(\lambda) = P^M(\lambda')$. Therefore, $P^M(\lambda)$ must display a discontinuity at $\lambda = 0$. This is the desired contradiction as we showed that $P^M(\cdot)$ is continuous. Thus, $P^M(\lambda') > P^M(0)$ must hold for all $\lambda' > 0$ and $\frac{dP^M(\lambda)}{d\lambda} > 0$ follows. □

A.2.2 Proof of Proposition 2

Proof. Note that this proposition holds for the *infinite* period model as well. Suppose it is optimal for the firm to choose $P^*(\widehat{\lambda}) = \widehat{P} \in \text{int}(\mathcal{P})$ for some $\widehat{\lambda} \in (0,1)$. We will show now that a firm's continuation value is equal to zero whenever it chooses its price equal to be $\widehat{P}$. Given some price P and prior belief λ_0, a firm's continuation value is defined as:

$$\beta\mathcal{V}(P;\lambda_0) \equiv \beta\bigg(\lambda_0\mathbb{E}_\varepsilon\left[V(b_1(\lambda_0,\log(P),\varepsilon))\right] + (1-\lambda_0)\mathbb{E}_\varepsilon\left[V(b_2(\lambda_0,\log(P),\varepsilon))\right]\bigg)$$

Recall that a firm faces a trade-off between maximizing current period expected profits and the value of information (through sharpening its posterior belief). The latter is captured by $\mathcal{V}(P;\lambda_0)$. As a result, a firm's marginal benefits are defined as:

$$\begin{aligned}
&\lambda_0\mathbb{E}_\varepsilon\left[V'(b_1(\lambda_0,\log(P),\varepsilon))\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial\log(P)}\frac{1}{P}\right]\\
+&(1-\lambda_0)\mathbb{E}_\varepsilon\left[V'(b_2(\lambda_0,\log(P),\varepsilon))\frac{\partial b_2(\lambda_0,\log(P),\varepsilon)}{\partial\log(P)}\frac{1}{P}\right]
\end{aligned}$$

It is straightforward to show that a firm's posterior belief at the *confounding* price $\widehat{P}$ equals its prior, i.e. we have:

$$\begin{aligned}
b_1(\lambda_0,\log(P),\varepsilon)\Big|_{P=\widehat{P}} &= b_2(\lambda_0,\log(P),\varepsilon)\Big|_{P=\widehat{P}}\\
&= \left(1+\frac{1-\lambda_0}{\lambda_0}\right)^{-1}\\
&= \lambda_0
\end{aligned}$$

for *all* $\varepsilon \in \mathbb{R}$. Also, the *expected* change in a firm's posterior belief at $P=\widehat{P}$ is exactly equal to zero as:

$$\begin{aligned}
\mathbb{E}_\varepsilon\left[\frac{\partial b_i(\lambda_0,\log(P),\varepsilon)}{\partial\log(P)}\bigg|_{P=\widehat{P}}\right] &= \mathbb{E}_\varepsilon\left[\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}(1-\lambda_0)\lambda_0(-1)^{\mathbf{1}(i=2)}\varepsilon\right]\\
&= 0
\end{aligned}$$

for $i \in \{1,2\}$ as $\mathbb{E}_\varepsilon[\varepsilon]=0$. Therefore, a firm's expected marginal benefit at $P=\widehat{P}$ reduces to:

$$V'(\lambda_0)\widehat{P}^{-1}\mathbb{E}_\varepsilon\left[\lambda_0\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial\log(P)}\bigg|_{P=\widehat{P}} + (1-\lambda_0)\frac{\partial b_2(\lambda_0,\log(P),\varepsilon)}{\partial\log(P)}\bigg|_{P=\widehat{P}}\right] = 0$$

If it is optimal for a firm to choose $P^*(\widehat{\lambda}) = \widehat{P}$, then it must be equal to $P^M(\widehat{\lambda})$ as there are no gains from experimentation. Recall that $P^M(0)=P_2^*$, $P^M(1)=P_1^*$ and $P^M(\cdot)$ is strictly increasing and continuous by proposition 1. Therefore, the confounding price $\widehat{P}\in\mathcal{P}$ is guaranteed to exist. Furthermore, proposition 1 and the *Intermediate Value Theorem* imply that there must exist some $\widehat{\lambda}$ such that $P^M(\widehat{\lambda})=\widehat{P}$.

By construction, we have $\widehat{P} \equiv \frac{\Delta\mu}{\Delta\sigma}$. Proposition 1 implies that $P^M(\cdot)$ is strictly increasing. As a result, it is straightforward to derive that the confounding belief is strictly increasing (decreasing) in $\Delta\mu$ ($\Delta\sigma$) as we must have $P^M(\widehat{\lambda}) = \frac{\Delta\mu}{\Delta\sigma}$. □

A.2.3 Proof of Proposition 3

Proof. We will show the case for $\lambda_0 = \frac{1}{2}$ in which $\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)} = -\frac{\partial b_2(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)}$ holds.[53] We want to derive sufficient conditions such that $P^*(\lambda_0) \in \text{int}(\mathcal{P})$ for all $\lambda_0 \in (0,1)$. This is equivalent to finding sufficient conditions such that a firm's expected marginal benefits strictly dominate its cost counterpart for $P = P_1^*$ and vice versa for $P = P_2^*$. More specifically, we need to show:

$$\eta\lambda_0\Pi_1'(P_2^*) + \frac{\beta}{P_2^*}\mathbb{E}_\varepsilon\Bigg[\bigg(\lambda_0 V_2'(b_1(\lambda_0,\log(P_2^*),\varepsilon)) - (1-\lambda_0)V_2'(b_2(\lambda_0,\log(P_2^*),\varepsilon))\bigg)\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)}\bigg|_{P=P_2^*}\Bigg] > 0 \qquad \textbf{(A1)}$$

$$(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \frac{\beta}{P_1^*}\mathbb{E}_\varepsilon\Bigg[\bigg(\lambda_0 V_2'(b_1(\lambda_0,\log(P_1^*),\varepsilon)) - (1-\lambda_0)V_2'(b_2(\lambda_0,\log(P_1^*),\varepsilon))\bigg)\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)}\bigg|_{P=P_1^*}\Bigg] < 0 \qquad \textbf{(A2)}$$

which is the first order condition with respect to P in period 1 evaluated at $P = P_2^*$ and $P = P_1^*$. We start by finding a sufficient condition for the first inequality **A1**. First, note that the sign of $\lambda_0 V_2'(b_1(\lambda_0,\log(P_2^*),\varepsilon)) - (1-\lambda_0)V_2'(b_2(\lambda_0,\log(P_2^*),\varepsilon))$ is completely determined by the sign of $b_1(\lambda_0,\log(P_2^*),\varepsilon) - b_2(\lambda_0,\log(P_2^*),\varepsilon)$ whenever $\lambda_0 = \frac{1}{2}$. This is true as $V_2(\cdot)$ is convex and therefore $V_2''(\cdot) > 0$.

It can be shown that $\exists\, \varepsilon(\lambda_0) < 0$ such that $b_1(\lambda_0,\log(P_2^*),\varepsilon) > b_2(\lambda_0,\log(P_2^*),\varepsilon)$ if and only if $\varepsilon < -\varepsilon(\lambda_0)$.[54] Define $x \equiv P_2^*\Delta\sigma - \Delta\mu < 0$, then we have that $\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)}\Big|_{P=P_2^*} > 0$ if and only if $\varepsilon > -x$. Now, denote $E_1 \equiv (-\infty, -\varepsilon(\lambda_0))$, $E_2 \equiv (-\varepsilon(\lambda_0), -x)$ and $E_3 \equiv (-x, +\infty)$. By construction, it must be that $E_1 \cup E_2 \cup E_3 = \mathbb{R}$.

The observations above imply that:

$$\bigg(\lambda_0 V_2'(b_1(\lambda_0,\log(P_2^*),\varepsilon)) - (1-\lambda_0)V_2'(b_2(\lambda_0,\log(P_2^*),\varepsilon))\bigg)\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)}\bigg|_{P=P_2^*} > 0$$

[53] The case for $\lambda_0 > \frac{1}{2}$ is extremely similar but less clean. We will include this case in a future draft of this paper.

[54] Observe that we have $-\varepsilon(\frac{1}{2}) = \frac{-x}{2}$, thus $-\varepsilon(\lambda_0) < -x$ whenever λ_0 is close enough to $\frac{1}{2}$.

for $\varepsilon \in E_2$. Thus, it is sufficient to show:

$$\eta\lambda_0\Pi_1'(P_2^*) + \frac{\beta}{P_2^*}\mathbb{E}_\varepsilon\Bigg[\Bigg(\lambda_0 V_2'(b_1(\lambda_0, \log(P_2^*), \varepsilon)) - (1-\lambda_0)V_2'(b_2(\lambda_0, \log(P_2^*), \varepsilon))\Bigg)\frac{\partial b_1(\lambda_0, \log(P), \varepsilon)}{\partial \log(P)}\Bigg|_{P=P_2^*}\Bigg|\varepsilon \in E_1 \cup E_3\Bigg] > 0$$

Observe the following strain of inequalities:

$$\begin{aligned}
&\eta\lambda_0\Pi_1'(P_2^*) + \tfrac{\beta}{P_2^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\big(x + \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \le -\varepsilon(\lambda_0)\right] - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \ge -x\right]\big) < \\
&\eta\lambda_0\Pi_1'(P_2^*) + \tfrac{\beta}{P_2^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\big(xF(-\varepsilon(\lambda_0)) + \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \le -\varepsilon(\lambda_0)\right] \\
&\qquad -x\left(1 - F(-x)\right) - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \ge -x\right]\big) = \\
&\eta\lambda_0\Pi_1'(P_2^*) + \tfrac{\beta}{P_2^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\left(\mathbb{E}_\varepsilon\left[x+\varepsilon|\varepsilon \in E_1\right] - \mathbb{E}_\varepsilon\left[x+\varepsilon|\varepsilon \in E_3\right]\right) \le \\
&\eta\lambda_0\Pi_1'(P_2^*) + \tfrac{\beta}{P_2^*}V_2'(1)\Bigg(\mathbb{E}_\varepsilon\left[\tfrac{\partial b_1(\lambda_0, \log(P), \varepsilon)}{\partial \log(p)}\Big|_{P=P_2^*}\ \Big|\ \varepsilon \in E_1\right] \\
&\qquad -\mathbb{E}_\varepsilon\left[\tfrac{\partial b_1(\lambda_0, \log(P), \varepsilon)}{\partial \log(p)}\Big|_{P=P_2^*}\ \Big|\ \varepsilon \in E_3\right]\Bigg) < \\
&\eta\lambda_0\Pi_1'(P_2^*) + \frac{\beta}{P_2^*}\mathbb{E}_\varepsilon\Bigg[\Bigg(\lambda_0 V_2'(b_1(\lambda_0, \log(P_2^*), \varepsilon)) - \\
&\qquad (1-\lambda_0)V_2'(b_2(\lambda_0, \log(P_2^*), \varepsilon))\Bigg)\frac{\partial b_1(\lambda_0, \log(P), \varepsilon)}{\partial \log(P)}\Bigg|_{P=P_2^*}\Bigg|\varepsilon \in E_1 \cup E_3\Bigg]
\end{aligned}$$

where the weak inequality follows from lemma 2 and the last strict inequality from the fact that $V_2'(1) > V_2'(\lambda')$ for any $\lambda' < 1$. This implies that we are done whenever we can show:

$$\eta\lambda_0\Pi_1'(P_2^*) + \tfrac{\beta}{P_2^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\big(x + \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \le -\varepsilon(\lambda_0)\right] - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \ge -x\right]\big) > 0$$

Recall that $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Therefore, we can use standard truncation formulas for our conditional expections. This gives us:

$$\begin{aligned}
\mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \le -\varepsilon(\lambda_0)\right] - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \ge -x\right] &= \frac{\varphi\left(\frac{-\varepsilon(\lambda_0)}{\sigma_\varepsilon}\right)}{\Phi\left(\frac{-\varepsilon(\lambda_0)}{\sigma_\varepsilon}\right)} - \frac{\varphi\left(\frac{-x}{\sigma_\varepsilon}\right)}{1-\Phi\left(\frac{-x}{\sigma_\varepsilon}\right)} \\
&> -\varphi(0)\left[\frac{1}{\Phi\left(\frac{-\varepsilon(\lambda_0)}{\sigma_\varepsilon}\right)} + \frac{1}{1-\Phi\left(\frac{-x}{\sigma_\varepsilon}\right)}\right] \\
&> -2\varphi(0)\left[\frac{1}{1-\Phi\left(\frac{-x}{\sigma_\varepsilon}\right)}\right]
\end{aligned}$$

Then, we can frame our first sufficient condition as:

$$\eta\lambda_0\Pi_1'(P_2^*) + \tfrac{\beta}{P_2^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\left[x - 2\frac{\varphi(0)}{1-\Phi\left(\frac{-x}{\sigma_\varepsilon}\right)}\right] \tag{\textbf{B1}}$$

In a similar fashion, we will derive a sufficient condition for **A2**. Once again, it can be shown that $\exists\, \varepsilon(\lambda_0) > 0$ such that $b_1(\lambda_0, \log(P_2^*), \varepsilon) > b_2(\lambda_0, \log(P_2^*), \varepsilon)$ if and only if $\varepsilon > -\varepsilon(\lambda_0)$. Let $y \equiv P_1^*\Delta\sigma - \Delta\mu > 0$, then for $\lambda_0 = \frac{1}{2}$, we have that $\varepsilon(\lambda_0) = \frac{y}{2} < y$. By straightforward algebra, it can be deduced that $\frac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial \log(P)}\Big|_{P=P_1^*} > 0$ if and only if $\varepsilon > -y$. Then, denote $E_1 \equiv (-\infty, -y)$, $E_2 \equiv (-y, -\varepsilon(\lambda_0))$ and $E_3 \equiv (-\varepsilon(\lambda_0), +\infty)$ which satisfies $E_1 \cup E_2 \cup E_3 = \mathbb{R}$.

This immediately implies that the following condition is sufficient for **A2** to hold:

$$\begin{aligned}(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \frac{\beta}{P_1^*}\mathbb{E}_\varepsilon\Bigg[\Bigg(&\lambda_0 V_2'(b_1(\lambda_0, \log(P_1^*), \varepsilon)) - \\ &(1-\lambda_0)V_2'(b_2(\lambda_0, \log(P_1^*), \varepsilon))\Bigg)\frac{\partial b_1(\lambda_0, \log(P), \varepsilon)}{\partial \log(P)}\Bigg|_{P=P_1^*}\Bigg|\varepsilon \in E_1 \cup E_3\Bigg] > 0\end{aligned}$$

We derive a similar chain of inequalities as before:

$$\begin{aligned}
&(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \tfrac{\beta}{P_1^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\big(y + \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \geq -\varepsilon(\lambda_0)\right] - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \leq -y\right]\big) < \\
&(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \tfrac{\beta}{P_1^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\big(y\left[1 - F(-\varepsilon(\lambda_0))\right] + \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \geq -\varepsilon(\lambda_0)\right] \\
&\qquad -yF(-y) - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \leq -y\right]\big) = \\
&(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \tfrac{\beta}{P_1^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\left(\mathbb{E}_\varepsilon\left[y + \varepsilon|\varepsilon \in E_3\right] - \mathbb{E}_\varepsilon\left[y + \varepsilon|\varepsilon \in E_1\right]\right) \leq \\
&(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \tfrac{\beta}{P_1^*}V_2'(1)\Bigg(\mathbb{E}_\varepsilon\left[\tfrac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial\log(p)}\Big|_{P=P_1^*}\ \Big|\ \varepsilon \in E_3\right] \\
&\qquad -\mathbb{E}_\varepsilon\left[\tfrac{\partial b_1(\lambda_0,\log(P),\varepsilon)}{\partial\log(p)}\Big|_{P=P_1^*}\ \Big|\ \varepsilon \in E_1\right]\Bigg) < \\
&(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \frac{\beta}{P_1^*}\mathbb{E}_\varepsilon\Bigg[\Bigg(\lambda_0 V_2'(b_1(\lambda_0, \log(P_1^*), \varepsilon)) - \\
&\qquad (1-\lambda_0)V_2'(b_2(\lambda_0, \log(P_1^*), \varepsilon))\Bigg)\frac{\partial b_1(\lambda_0, \log(P), \varepsilon)}{\partial \log(P)}\Bigg|_{P=P_1^*}\Bigg|\varepsilon \in E_1 \cup E_3\Bigg]
\end{aligned}$$

where the weak inequality follows from lemma 2 and the last strict inequality from the fact that $V_2'(1) > V_2'(\lambda')$ for any $\lambda' < 1$. This implies that we are done whenever we can show:

$$(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \tfrac{\beta}{P_1^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\big(y + \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \geq -\varepsilon(\lambda_0)\right] - \mathbb{E}_\varepsilon\left[\varepsilon|\varepsilon \leq -y\right]\big) < 0$$

Using the previous finding on expectations of truncated standard normal random variables, the latter inequality is satisfied whenever the following condition holds:

$$(1-\eta)(1-\lambda_0)\Pi_2'(P_1^*) + \tfrac{\beta}{P_1^*}V_2'(1)\tfrac{\Delta\sigma}{\sigma_\varepsilon^2}\left(y + 2\frac{\varphi(0)}{\Phi\left(\frac{-y}{\sigma_\varepsilon}\right)}\right) < 0 \tag{\textbf{B2}}$$

Whenever we define $x_{\min}$ and $x_{\max}$ as:

$$x_{\min} \equiv \log(P_2^*)\Delta\sigma - \Delta\mu - \frac{2\varphi(0)}{1-\Phi\left(-\frac{\log(P_2^*)\Delta\sigma-\Delta\mu}{\sigma_\varepsilon}\right)} < 0,$$
$$x_{\max} \equiv \log(P_1^*)\Delta\sigma - \Delta\mu + \frac{2\varphi(0)}{\Phi\left(-\frac{\log(P_1^*)\Delta\sigma-\Delta\mu}{\sigma_\varepsilon}\right)} > 0.$$

then, it is clear that **B1** and **B2** are satisfied whenever $V_2'(1)$ is bounded from above. More precisely, we get:

$$V_2'(1) < \overline{\mathcal{V}} \equiv \frac{\sigma_\varepsilon^2}{\beta\Delta\sigma}\min\left\{\frac{\eta\lambda_0\Pi_1'(P_2^*)}{-x_{\min}}, \frac{(1-\eta)(1-\lambda_0)\left(-\Pi_2'(P_1^*)\right)}{x_{\max}}\right\}. \tag{\textbf{B}}$$

Thus, we have shown **B** $\Longrightarrow$ (**B1** and **B2**) $\Longrightarrow$ (**A1** and **A2**). However, we concluded in the beginning of the proposition that $P^*(\lambda_0) \in \text{int}(\mathcal{P})$ whenever **A1** and **A2** hold. This is exactly what we wanted to show. $\square$

A.3 Numerical Algorithm

Parameters. β, δ, σ_1, σ_2, λ_0, m, s, σ_ε, ρ and σ_ζ.
Numéraire. $W = 1$.

Algorithm: pseudo-code.

We assume consumer taste shocks to satisfy $\varepsilon_k \sim N(m, s^2)$. As a result, integrals over consumer taste shocks are approximated using Gaussian quadrature methods. With some abuse of notation, let the weights and nodes be denoted by $\{\omega_j^{GH}\}_{j=1}^M$ and $\{\zeta_j^{GH}\}_{j=1}^M$ respectively.[55] The quadrature weights and nodes are chosen "optimally". The nodes $\{\zeta_j^{GH}\}_{j=1}^M$ are the *roots* of the Hermite polynomial $H_M(\zeta)$ which is defined as $H_M(\zeta) = \oint \frac{M!}{2\pi i}\exp\left(-t^2+2t\zeta\right)t^{-(M+1)}dt$ and the weights are equal to:

$$\omega_i^{GH} = \tfrac{2^{M-1}M!\sqrt{\pi}}{M^2 H_{M-1}(\zeta_j^{GH})^2}$$

We approximate the continuous AR(1) process for idiosyncratic productivity with a finite state Markov process by following the Tauchen (1986) procedure. The number of Markov states will be

[55] The superscript stands for "Gaussian-Hermite" quadrature. This is useful to approximate functions of the form $f(x) = \exp(-x^2)$ which includes the family of normal distributions.

denoted by N_T. Let the Markov transition density be denoted by $\mathcal{M}(z_i, z_j)$ for $i, j \in \{1, 2, \ldots, N_T\} \times \{1, 2, \ldots, N_T\}$.

I (initialization). Set $\overline{P}_1^0$, $\overline{P}_2^0$, χ^0 and convergence criteria $\Delta_\varepsilon, \Delta_\varphi > 0$. Let the counter k equal 0.

II (out). Given k, set $\overline{P}_1^k$, $\overline{P}_2^k$ and χ^k.

III (in). Set $H = \frac{1}{3}$ and calculate Π^k using:

$$H^\eta = \frac{1}{\chi^k} \frac{1}{H + \Pi^k}$$

Then, set $S^k = H + \Pi^k$. Define the aggregate state as $\omega^k = (\overline{P}_1^k, \overline{P}_2^k, S^k)$.

IV. Solve the firm's problem by obtaining $V(\lambda, z, p_{-1})$:

$$\begin{aligned}
V(\lambda, z, p_{-1}) &= \max\left\{V^A(\lambda, z), V^N(\lambda, z, p_{-1})\right\} \text{ where} \\
V^A(\lambda, z) &= \max_{p \geq 0} (p - \tfrac{1}{z}) \left[\lambda\eta \tfrac{p^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda)(1-\eta) \tfrac{p^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}}\right] \tfrac{1}{\overline{P}_1^\eta \overline{P}_2^{1-\eta} \chi} - \tfrac{\psi}{\overline{P}_1^\eta \overline{P}_2^{1-\eta}} \\
&\quad + \beta\lambda \sum_{i=1}^{N_T} \sum_{j=1}^{M} \mathcal{M}(z_i, z) \tfrac{1}{\sqrt{\pi}} \omega_j^{GH} V\left(b_1(\lambda, \log(\tfrac{p}{1+\tilde{\pi}}), \sqrt{2}\sigma_\alpha \zeta_j^{GH} + \mu_\alpha\right), z_i, \tfrac{p}{1+\tilde{\pi}}) \\
&\quad + \beta(1-\lambda) \sum_{i=1}^{N_T} \sum_{j=1}^{M} \mathcal{M}(z_i, z) \tfrac{1}{\sqrt{\pi}} \omega_j^{GH} V\left(b_2(\lambda, \log(\tfrac{p}{1+\tilde{\pi}}), \sqrt{2}\sigma_\alpha \zeta_j^{GH} + \mu_\alpha\right), z_i, \tfrac{p}{1+\tilde{\pi}}) \\
V^N(\lambda, z, p_{-1}) &= (p_{-1} - \tfrac{1}{z}) \left[\lambda\eta \tfrac{p_{-1}^{-\sigma_1}}{\overline{P}_1^{1-\sigma_1}} + (1-\lambda)(1-\eta) \tfrac{p_{-1}^{-\sigma_2}}{\overline{P}_2^{1-\sigma_2}}\right] \tfrac{1}{\overline{P}_1^\eta \overline{P}_2^{1-\eta} \chi} \\
&\quad + \beta\lambda \sum_{i=1}^{N_T} \sum_{j=1}^{M} \mathcal{M}(z_i, z) \tfrac{1}{\sqrt{\pi}} \omega_j^{GH} V\left(b_1(\lambda, \log(\tfrac{p_{-1}}{1+\tilde{\pi}}), \sqrt{2}\sigma_\alpha \zeta_j^{GH} + \mu_\alpha\right), z_i, \tfrac{p_{-1}}{1+\tilde{\pi}}) \\
&\quad + \beta(1-\lambda) \sum_{i=1}^{N_T} \sum_{j=1}^{M} \mathcal{M}(z_i, z) \tfrac{1}{\sqrt{\pi}} \omega_j^{GH} V\left(b_2(\lambda, \log(\tfrac{p_{-1}}{1+\tilde{\pi}}), \sqrt{2}\sigma_\alpha \zeta_j^{GH} + \mu_\alpha\right), z_i, \tfrac{p_{-1}}{1+\tilde{\pi}}) \\
&\text{with } b_i(\lambda, \log(p), \epsilon) = \left[1 + \frac{1-\lambda}{\lambda} \frac{F'(\mu_i - \mu_2 + (\sigma_2 - \sigma_i)\log(p) + \epsilon)}{F'(\mu_i - \mu_1 + (\sigma_1 - \sigma_i)\log(p) + \epsilon)}\right]^{-1} \\
&\text{and } \mu_i = (\sigma_i - 1)\log(\overline{P}_i) + \log(\eta_i)
\end{aligned}$$

IV. Store the optimal pricing policy function $P^*(\lambda, z)$ for every $(\lambda, z) \in [0, 1] \times Z$.

V. Simulation. Simulate a panel of $N = 50000$ firms who use the policy function $P^*(\lambda, z)$.

Simulation initialization. The initial distribution $\varphi_{i,0}(\lambda, z)$ for $i = 1, 2$ is degenerate at (λ_0, z_0). For each firm $n \in \{1, 2, \ldots, N\}$, assign it to be a firm of type $\sigma_n = \sigma_1$ with probability λ_0. Set time counter t to zero.

V.a. Given a firm's belief $\lambda_{n,t}$, let firm n set price $P^*(\lambda_{n,t}, z_n)$. Generate log sales by drawing log demand shocks $\varepsilon_{n,t} \sim N(m, s^2)$ through:

$$q_{n,t} = -\sigma_n P^*(\lambda_{n,t}, z_n) + \mu_i + s^k + \varepsilon_{n,t}$$

where $\mu_i = (\sigma_i - 1)\log(\overline{P}_i^k) + \log(\zeta_i)$. Update the firm n's posterior to:

$$\lambda_{n,t+1} = B(\lambda_{n,t}, P^*(\lambda_{n,t}, z_n), q_{n,t}, S^k)$$

Apply exogenous death shocks δ for each firm. If a firm exits, then replace it by a new firm which is assigned to be a type σ_1 firm with probability λ_0. Its prior becomes λ_0.

V.b. Calculate $\varphi_{i,t+1}(\lambda, z)$ for each $i = 1, 2$. Stop the simulation when the distribution of beliefs settles in both measures of active firms or when the number of simulation periods exceed some upper bound $T > 1$, i.e. $\sup_{\lambda \in (0,1), z \in Z} ||\varphi_{i,t+1}(\lambda, z) - \varphi_{i,t}(\lambda, z)|| < \Delta_\varphi$ for $i \in \{1, 2\}$ and/or $t = T$. Otherwise, set $t := t + 1$ and repeat step **V.a**.

VII. Calculate $\overline{P}_i^{temp}$ with the *simulated* density $\tilde{\Phi}_i(\lambda)$:

$$\overline{P}_i^{temp} = \left(\sum_{\lambda} P^*(\lambda, z)^{1-\sigma_i} \tilde{\Phi}_i(\lambda, z) \right)^{\frac{1}{1-\sigma_i}}$$

where $\tilde{\Phi}_i(\lambda, z)$ is the *empirical* cross-sectional probability distribution function of beliefs and idiosyncratic productivity. Also, calculate the total amount of labor in the economy as:

$$H^{temp} = S^k \sum_{i=1}^{2} \frac{\eta_i}{z} \left[\frac{\sum_{\lambda,z} P^*(\lambda, z)^{-\sigma_i} \tilde{\Phi}_i(\lambda, z)}{\sum_{\lambda,z} P^*(\lambda, z)^{1-\sigma_i} \tilde{\Phi}_i(\lambda, z)} \right]$$

If $\sup_i |\overline{P}_i^{temp} - \overline{P}_i^k| < \Delta_\varepsilon$ **and** $|H - H^{temp}| < \Delta_\varepsilon$, then stop; otherwise, set $\overline{P}_1^{k+1} = \overline{P}_i^{temp}$ and $\overline{P}_2^{k+1} = \overline{P}_2^{temp}$. Let $\chi^{k+1} > \chi^k$ if and only if $H - H^{temp} < 0$. Update the counter to $k := k + 1$ and repeat step **II**.

A.4 Empirical Supplement

A.4.1 Identifying New Products

We assume that if a UPC changes, it is likely that some noticeable characteristic of the product has also change. This is because it is rare that a meaningful quality change occurs without resulting in a UPC change. The same assumption was made in Broda and Weinstein (2010). Considering each UPC as a product is, in fact, a very broad definition since it includes classically innovative products, line extensions (including quality upgrades), and temporary items (i.e. seasonal). We find that product line extensions, such as flavor or form upgrades or novelty/seasonal items, are much more prevalent than new brand introduction. Finally, to minimize the problem that some UPCs might get discontinued only to have the same product appear with a new UPC, as noted by Chevalier et al. (2003), we only consider products that lasted at least two years in the market.

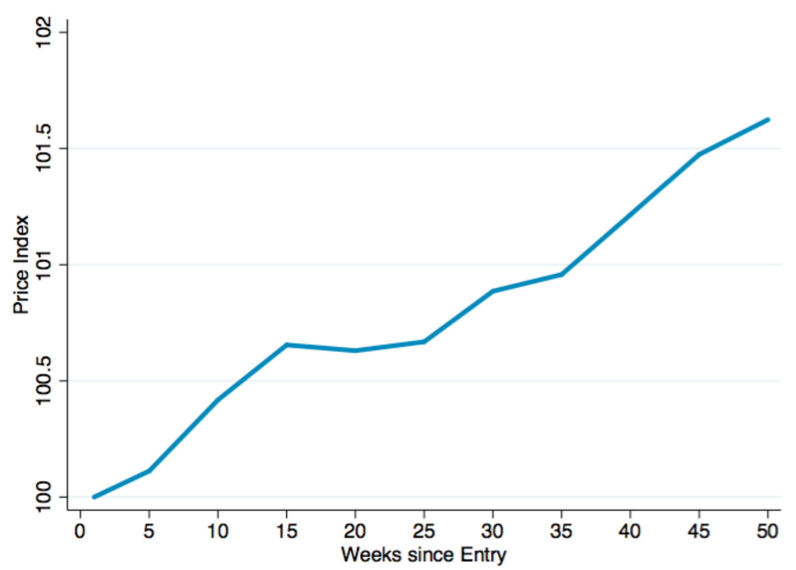
Using the UPC identifier and the retailer identifier provided by the data, we determine the specific week and store in which each product first appeared. We consider as entering products those that enter the market after January 2002. Since our data starts from January 2001, an entering product is one that had no observable transactions, in any store across the U.S., for at least one year. This assumption allows us to avoid including products with left censored age.

Finally, since IRI Symphony used different hierarchical assignments for UPCs starting on 2007, the data includes some entering and exiting UPCs that might not correspond to new product introductions. IRI used different product stubs for 2001-2006, 2007, and 2008-2011. A stub is the hierarchical assignment in the data (i.e. UPCs to brands, brands to vendors, vendors to types of categories). For example, vendor and parent could change due to merger and acquisition activity, brand name could have changed from 2006 to 2007, etc.

In addition, for each different product stub, IRI undertook the following actions: i) reorganized private-label items (i.e. in some categories they break out organic private label), ii) dropped UPCs that have not moved in the past years, iii) collapsed UPCs into a main UPC to avoid clutter (i.e. products that came to a store as part of a special promotional code rather than with a standard UPC code), iv) reorganized categories (i.e. a category migh have increased in scope and as a consequence suffered an increase in items) and iv) added UPCs that were introduced at the beginning of each stub. All of these are consistent with changes in the number of entering and exiting UPCs due to changes in the product stub rather than new product introductions or products being phased out. To avoid this problem, we only consider products that entered before the first week of 2007.

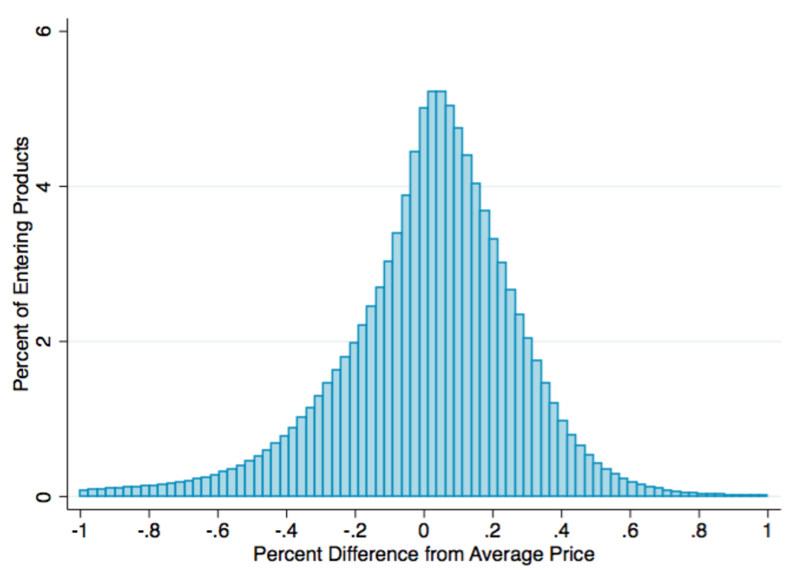
A.4.2 Price Increases and Decreases at Entry

Figure 10: Price Index for New Products



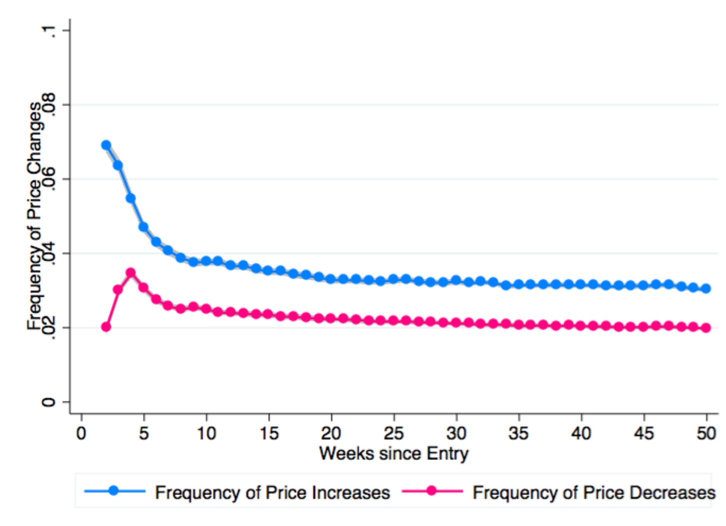
Note: The graph plots a geometric price index for new products. It considers the first year since entry. The expenditure weights used are at the UPC level and based on the first year of sales of each product. The data source is the Symphony IRI dataset.

Figure 11: Distribution of Entry Prices



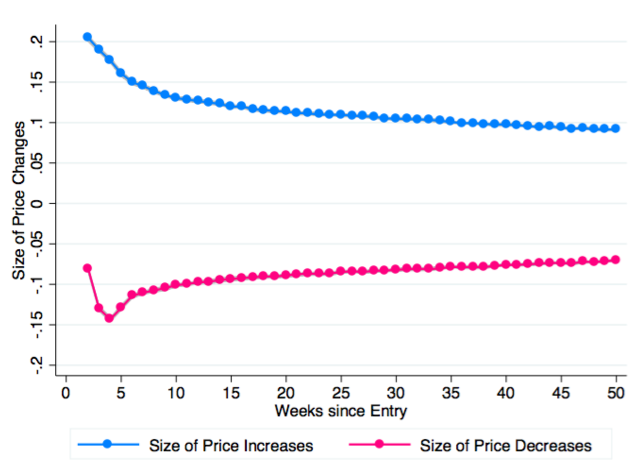
Note: The graph plots the percent difference between the entry price of all new products in our sample with respect to products of the same size, within the same category, at the store in which they were launched. The data source is the Symphony IRI dataset.

Figure 12: Frequency of Price Increases and Decreases at Entry



Note: The graph plots the average weekly frequency of price adjustments of products entering the market. The y-axis denotes the probability that the product adjusts prices in a given week and the x-axis denotes the number of weeks the product has been observed in the data since it entered the market. The graph plots the age fixed effects where we used a regular price change indicator as dependent variable controlling for store, UPC, and time fixed effects. The blue line indicates the frequency of positive price adjustments and the red line the frequency of negative price adjustments. The calculation uses approximately 130 million observations and 2.5 million stores$\times$UPC pairs. The data source is the Symphony IRI dataset.

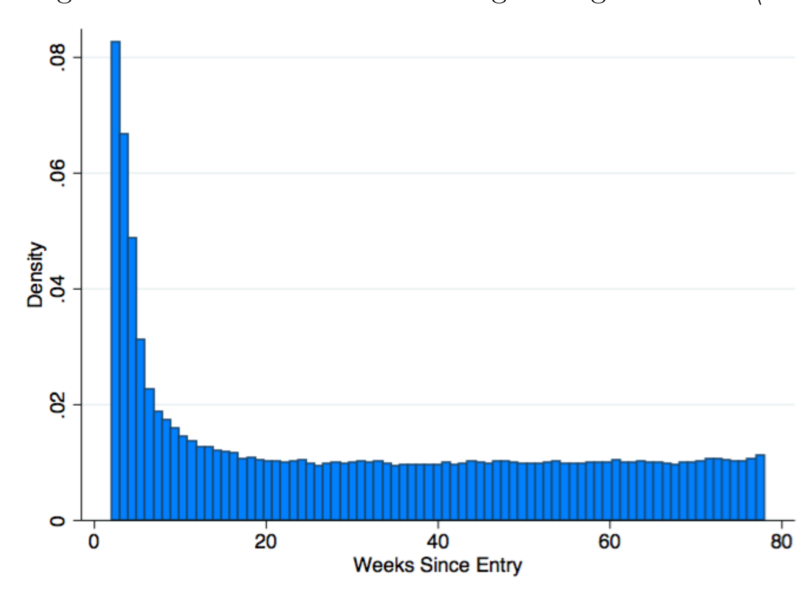
Figure 13: Absolute Value of Price Increases and Decreases at Entry



Note: The graph plots the average size of price adjustments of products entering the market. The y-axis is the value of the log price change in that week and the x-axis denotes the number of weeks the product has been observed in the data since it entered the market. The graph plots the age fixed effects where we used the log price change as dependent variable controlling for store, UPC, and time fixed effects. The blue line indicates the average size of positive price adjustments and the red line the average size of negative price adjustments. The calculation uses approximately 5.8 million price changes and 2.5 million stores$\times$UPC pairs. The data source is the Symphony IRI dataset.

A.4.3 Large Price Changes (Robustness)

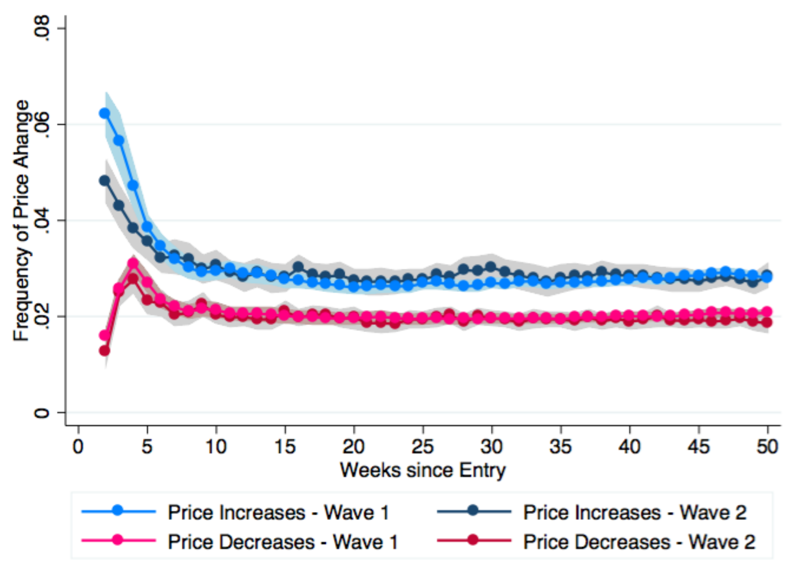
Figure 14: Fraction of Price Changes Larger than 30\%



Note: The figure shows the fraction of price changes larger than 30% in a given category and city as a function of the age of the product. The products considered are those that last at least two years in the market. Source: IRI Symphony dataset

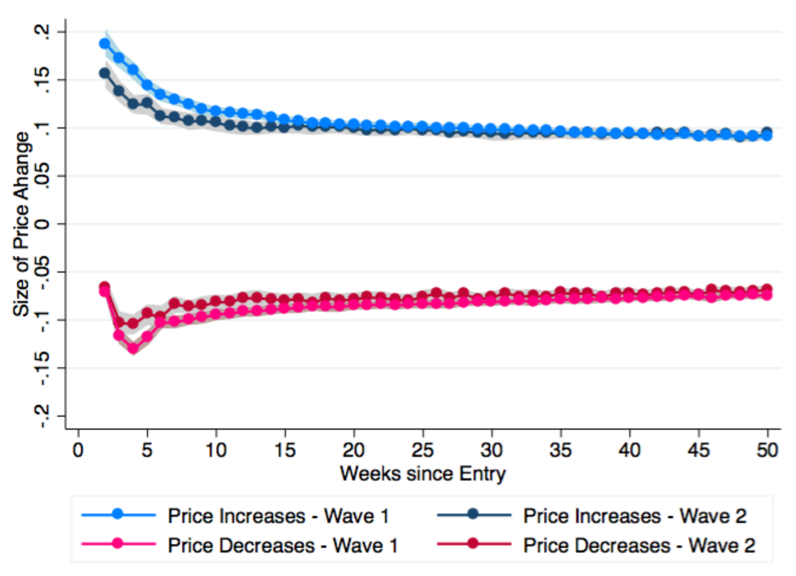
A.4.4 Price Increases and Decreases by Waves

Figure 15: Frequency of Price Adjustment (Positive and Negative)



Note: The figure shows the probability of price adjustment with respect to the mean for both price increases and price decreases. Wave 1 represents products that were launched during the first year since the product was introduced. Wave 2 represents the same products when launched in different stores a year later. The graphs shown control for stores, time and products fixed effects.

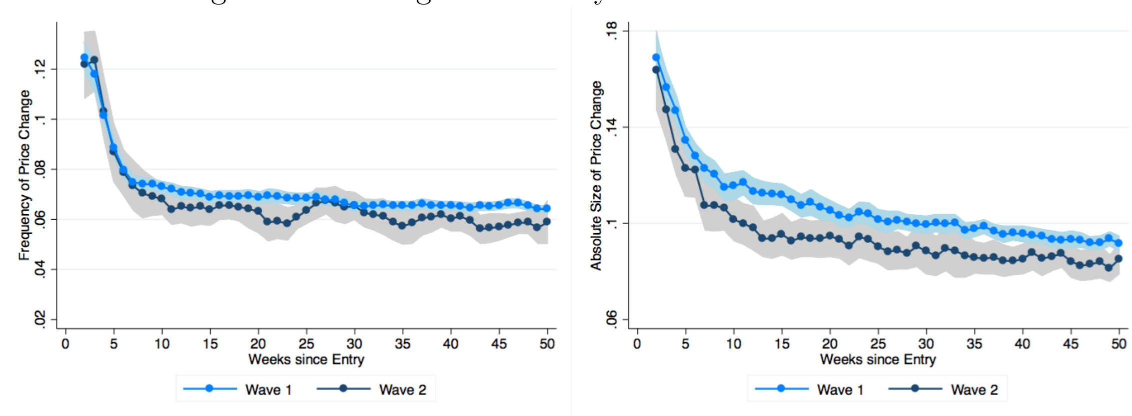
Figure 16: Size of Price Adjustments (Positive and Negative)



Note: The figure shows the size of prices changes with respect to the mean for both price increases and price decreases. Wave 1 represents products that were launched during the first year since the product was introduced. Wave 2 represents the same products when launched in different stores a year later. The graphs shown control for stores, time and products fixed effects.

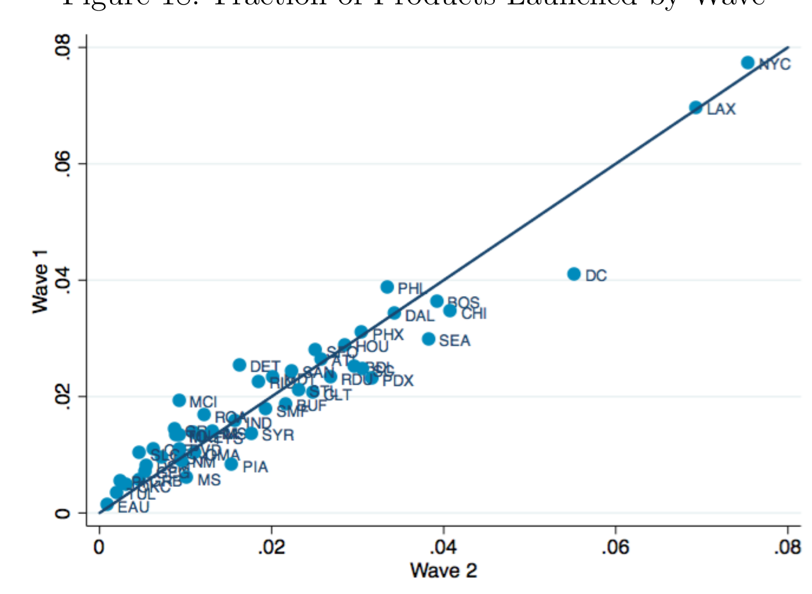
Pricing Moments by Waves (Different City)

Figure 17: Pricing Moments by Waves in Different Cities



Note: The figure shows the probability of adjusting prices and the size of adjustment by waves. Wave 1 represents products that were launched during the first year since the product was introduced. Wave 2 represents the same products when launched in different stores (located in different cities) a year later. The graphs shown control for stores, time and products fixed effects.

Figure 18: Fraction of Products Launched by Wave



Note: The figure shows the fraction of products launched in each wave by MSA.

A.4.5 Entry and Exit Rates at the Product Level

We distinguish between products entering or exiting the market and products being launched or phased out at each store, where our unit of observation is every UPC×Store pair. We document a substantial degree of entry and exit of products at both levels.[56] This fact was first documented by Broda and Weinstein (2010) who use data directly collected from consumers to document that on average 10 percent of household expenditures are on goods that were created in the last year.[57] Because their dataset only allows them to observe products households in their sample purchased, the entry rates that they calculate can be seen as the rate products are adopted or discontinued by households. Using a scanner dataset collected at the store level offers the advantage of observing, for the categories available, the entire universe of products for which a transaction is recorded in a given week. This allow us to get a closer approximation to the actual rate of entry and exit of products in the market.

We define the entry and exit rates, as well as the rate of creation and destruction, in the same way as Broda and Weinstein (2010). The entry rate is defined as the number of new products in period t relative to period s as a share of the total number of products purchased in period t. A new product is one that records at least one transaction in period t in any of the stores in our sample and that was not sold in any store in period s. The exit rate is defined in a similar way.

$$\textit{Entry Rate}(t,s) = \frac{\#\,\textit{New}\,UPCs(t,s)}{\#\,\textit{All}\,UPCs(t)}$$

$$\textit{Exit Rate}(t,s) = \frac{\#\,\textit{Disappearing}\,UPCs(t,s)}{\#\,\textit{All}\,UPCs(s)}$$

Creation and destruction are the revenue weigthed analogues of the entry and exit rates. As a result, net creation can be defined as the difference between creation and destruction and product turnover as the sum. Table III reports the entry and exit rates for the case in which t and s are one and five years apart.

The table shows that 15 percent of the UPCs in the market and on average 27 percent of the products in each store entered in the last year. In addition, approximately 45% of the products in the market entered in the last five years accounting for 30% of total expenditures. At the store level, 66% of all products sold were first introduced by the store in the last five years and they account for more than half of the total revenue of the store. Although the exit rate is very similar to the entry rate of products, destruction is lower than creation, indicating that consumers spend more on new products than on products about to exit.

[56] In order to avoid including products with right censored age, a product exiting the market must have its last transaction recorded before the last week of 2010, which is one year before the last week in our datset. In addition, we only consider products that did not exit exactly the weeks when the new product stub was implemented.

[57] Both Bernard et al. (2010) and Broda and Weinstein (2010) document that more than 90 percent of the product creation observed in the data occurs within existing firms.

Table III: Product Entry and Exit

	UPC	UPC	UPC×Store	UPC×Store
Period	5-Year	1-Year (mean)	5-year	1-year (mean)
Entry	0.45	0.14	0.66	0.27
Creation	0.29	0.07	0.47	0.15
Exit	0.42	0.13	0.61	0.25
Destruction	0.08	0.01	0.39	0.10

Note: Entry rate = Number of new UPCs(t) / Total number of UPCs(t), Creation = Value of new UPCs(t) / Total value(t), Exit rate = Number of disappearing UPCs(t-1) / Total number of UPCs(t-1), Destruction = Value of disappearing UPCs(t-1) / Total value (t-1).

The rate of product turnover indicates that, at any point in time, there is a large amount of products being launched or being phased out at each point in time. Panel A in table IV shows that the median duration of a UPC×Store pair is only around two years.[58] Panel B shows the distribution of durations counting the number of observations (weeks with at least one transaction) since entry. The median UPC×Store pair is sold on approximately 47 different weeks on average.

Table IV: Distribution of Duration by UPCxStore

Panel A: Weeks since Entry

p(1)	p(25)	p(50)	p(75)	p(99)	Mean	Std
1.03	37.41	96.35	209.54	450.77	134.03	118.94

Panel B: Observations since Entry

p(1)	p(25)	p(50)	p(75)	p(99)	Mean	Std
1.03	16.32	47.12	122.04	369.74	83.12	90.67

Note: The table shows the descriptive statistics of the distribution of duration of a UPC×Store pair. We computed the duration of each UPC×Store pair and aggregated them to the category level using equal weights. Categories were further aggregated using equal weights. Panel A shows the statistics for the number of weeks since the products entered. Panel B shows the statistics for the number of times a product is observed in our data set. A product is observed only if it records a transaction in a given week and store.

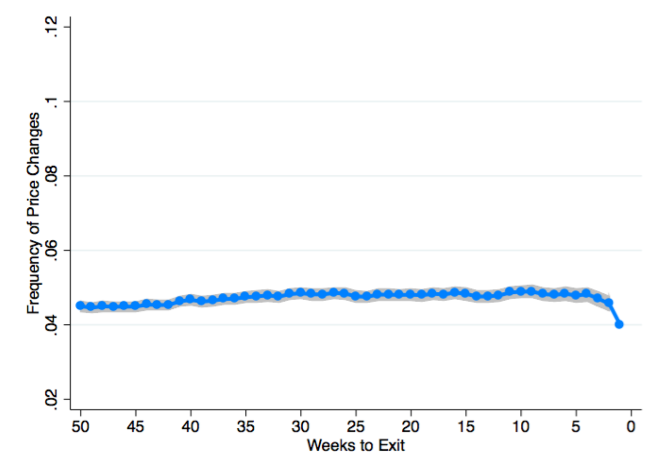
A.4.6 Product Life Cycle

Exit

Figure 19 shows that the frequency of price changes stays mostly constant and decreases only around 1 percentage point near exit. This is the case for both price increases and decreases. Similarly, figure 20 shows that the absolute value of price changes stays close to its average value (around 10%) during the last weeks of the product existence in the market.

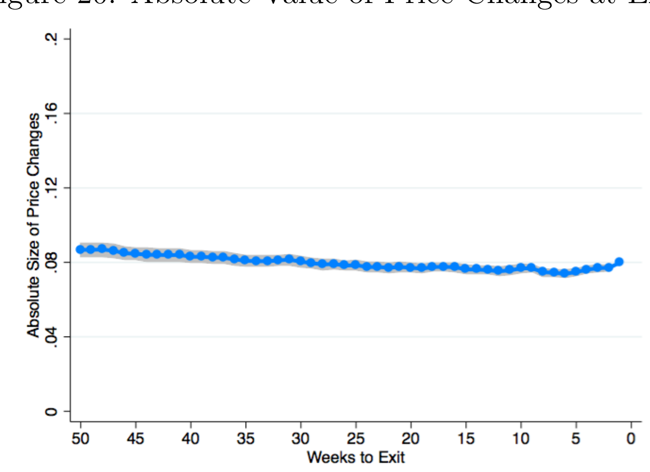
[58] Since our dataset ends the last week of 2011 and we are considering products that entered the last week of 2006 at the latest, right censoring is only an issue for products that last more than 261 weeks in the market.

Figure 19: Frequency of Price Adjustment at Exit



Note: The graph plots the average weekly frequency of price adjustments of products exiting the market. The y-axis denotes the probability that the product adjusts prices in a given week and the x-axis denotes the number of weeks a product has left in the market before exiting. The graph plots the age fixed effects coefficients of regression where we used the regular price change indicator as dependent variable and we control for store, UPC, and time fixed effects. The calculation uses approximately 130 million observations and 2.5 million stores×UPC pairs. The standard errors are clustered at the store level. The data source is the Symphony IRI dataset.

Figure 20: Absolute Value of Price Changes at Exit



Note: The graph plots the average absolute size of price adjustments of products entering the market. The y-axis is the absolute value of the log price change in that week and the x-axis denotes the number of weeks a product has left in the market before exiting. The graph plots the age fixed effects coefficients of regression where we used the absolute value of the log price change as dependent variable and we control for store, UPC, and time fixed effects. The calculation uses approximately 5.8 million price changes and 2.5 million stores×UPC pairs. The standard errors are clustered at the store level. The data source is the Symphony IRI dataset.