Coarse Pricing Policies*

Luminita Stevens

University of Maryland

June 2019

Abstract

The muted volatility of inflation during the Great Recession and its aftermath has refocused attention on the constraints that firms face when adjusting prices. Using new empirical and theoretical results, I argue that each firm's *choice* of how much information to acquire to set prices plays a central role in determining the patterns of pricing at the product-level and the degree of aggregate price rigidity in response to shocks. In support of the information channel, I present product-level evidence that firms price goods using coarse pricing policies that are updated infrequently and consist of a small menu of prices. Firms are heterogeneous in the complexity and duration of their pricing policies, and this heterogeneity is reflected in differential responses to the Great Recession cycle, with firms exhibiting more complex policies responding more aggressively. I develop a theory of information-constrained price setting that generates coarse pricing endogenously, and quantitatively matches the discreteness, duration, and volatility of policies in the data. The information friction dampens the responsiveness of prices to shocks, and, coupled with heightened volatility, induces firms to keep prices relatively high, to protect against losses in an uncertain environment.

*I am grateful to the editor and three referees for their very valuable feedback, and to Mike Woodford and Ricardo Reis for their advice and support. I would also like to thank and Boragan Aruoba, Ryan Chahrour, Doireann Fitzgerald, Christian Hellwig, Pat Kehoe, Ben Malin, Filip Matějka, Virgiliu Midrigan, Emi Nakamura, Jaromir Nosal, Ernesto Pasten, Chris Sims, Jón Steinsson, and participants at numerous seminars and conferences for helpful comments. Camilo Morales-Jimenez, Dun Jia, and Donggyu Lee provided excellent research assistance. This paper supersedes the earlier working papers Stevens (2011; 2012).
Contact: stevens@econ.umd.edu.

1 Introduction

The behavior of inflation during the Great Recession and its aftermath has challenged conventional models of price adjustment. First, the United States experienced an unexpectedly mild disinflation during the most severe downturn since the Great Depression. Second, inflation was slow to pick up during the subsequent recovery.[1] What accounts for these muted inflation dynamics in the midst of such turbulence in economic activity remains an open question. Using data and theory, I argue that information frictions—specifically firms' *choices* of how much information to acquire to set prices—played a key role in shaping product-level pricing patterns and in dampening aggregate price dynamics during this period.

In support of the mechanism of endogenous information frictions, I first present evidence that firms set prices using plans that are sticky and coarse. I identify these plans by searching for changes in the distribution of prices charged over time for each individual good. I detect these change points using an adaptation of the Kolmogorov-Smirnov test, which allows for any change in either the shape or the support of a distribution. Applied to weekly scanner price data covering the 2006 to 2015 period, the method identifies pricing policies that change every seven to eight months, and typically consist of a menu of three to four distinct price points, among which the firm alternates roughly every three weeks.

The discreteness of price levels despite the high frequency of price changes suggests that while the *timing* of price adjustment is quite flexible, the *level* to which the price adjusts is more constrained. This pattern is at odds with prior models of rigid prices, which assume that the timing of adjustment is constrained—exogenously or due to menu costs—but that once the firm decides to adjust, a new price is chosen optimally.[2] As I show in the second part of the paper, this pattern arises endogenously in a model of information-constrained pricing, as a cheap way for firms to crudely track the optimal full information price.

As is well-known in the literature, there are large differences across products in the frequency of price changes (Nakamura & Steinsson, 2008). I provide an alternative classification of products, in terms of the types of pricing policies employed. I identify three broad types of policies. Approximately 12% of products feature *single price policies* (SPP), like the canonical time-dependent or state-dependent pricing models. These products adjust their prices much less frequently and by smaller amounts than average, suggesting that they face a relatively low volatility of their target price. On the other hand, 60% of products exhibit policies with *multiple rigid prices* (MRP). The volatility of the data is concentrated in these products, which display very frequent and large price changes. However, despite

[1] Hall (2011), Ball & Mazumder (2011), Watson (2014), and Del Negro, Giannoni & Schorfheide (2015).
[2] An exception is the theory of rigid pricing due to ambiguity aversion by Ilut, Valchev & Vincent (2019).

this volatility, only three-to-four distinct price levels are typically realized over the life of a policy realization. These products seem to face a high volatility in their desired price, to which they respond by picking a small set of prices among which to alternate, and then occasionally updating this set. The coarseness and volatility of these product prices pose the biggest challenge to existing pricing models, but can be rationalized as a way for firms to economize on information costs. The remaining 28% of products are characterized by *one-to-flex policies* (OFP), in which one rigid price is accompanied by flexible, short-lived deviations to and from it. This type of pattern has been generated in models that endow firms with different technologies for setting regular versus temporary prices (Kehoe & Midrigan, 2015; Guimaraes & Sheedy, 2011). In the data, these goods feature large and frequent policy shifts, but muted within-policy price volatility. They seem to face more volatility in their desired price than the SPP goods, but also relatively high costs of implementing more complex policies. The theory proposed generates this range of policy types endogenously, as a function of differences in fundamental parameters.

Classifying products by policy type proves useful for understanding inflation dynamics during the Great Recession. Inflation differs significantly across policy types in terms of its volatility and its sensitivity to the state of the economy. While the inflation rates for all product types moved in tandem during the relatively tranquil periods at the beginning and end of the sample, they diverged substantially during the recession and its immediate aftermath. Once inflation started to fall in late 2008, it fell twice as much for the MRP goods as for the SPP goods. Moreover, SPP products continued to raise prices throughout the crisis, while MRP products actually cut prices. During this period, the SPP inflation rate was much less volatile, while the MRP inflation responded much more aggressively to the cycle. The information-based theory presented in this paper predicts precisely these effects, through the information acquisition channel: Firms that generally operate in more volatile markets have incentives to acquire more accurate information, which in turn enables them to choose more complex pricing policies; as a result, they also respond to the aggregate state of the economy more aggressively.

These findings underscore the value of studying pricing data in its entirety, rather than filtering out temporary price changes. Transitory price volatility is crucial to identifying the type of policy of each good, and moreover, the dynamics of different policies during the Great Recession show that transitory price volatility is at least partially responsive to the aggregate state, and does not wash out in the aggregate. This result qualifies the pricing literature's conclusion that micro price volatility is not relevant to aggregate rigidity. The proposed theory then quantifies the magnitude of this effect.

Distinguishing between policy changes and raw price changes is also useful for evaluating alternative theories of price setting and identifying potential sources of shocks. The dynamics of price and policy adjustment over time illustrate this point. First, the frequency and size of within-policy price changes are positively correlated over the sample period. This correlation is difficult to reconcile with models of menu costs, which would predict a negative correlation. Instead, it suggests heterogeneity in the volatility of market conditions faced by different firms: The prices of some products rarely change, and even when they do, they change by modest amounts, while others feature the opposite pattern. Second, the rate of policy adjustments rose during the Great Recession, suggesting at least partial state-dependence in the updating of policies over time, ruling out Calvo-like policy adjustment. At the same time, neither the rate nor the size of raw price changes differed significantly. Moreover, the incidence of multi-rigid price policies declined in the recession, while the incidence of single-price policies rose. These patterns point to the role of heightened volatility in shaping price dynamics during this decade. Firms responded to the increase in volatility associated with the Great Recession not by making their pricing plans more complex, but rather by keeping them simple and reviewing them often. This interpretation is further supported by the increase in the rate of policy changes and in the incidence of single price policies that occurred in 2011, which was another period of increased volatility. The theory of information frictions presented in the second part of the paper generates these patterns of policy adjustment in response to heightened volatility.

What drives the large within-policy price volatility? How much of it reflects responding to shocks versus price discriminating (PD) among heterogeneous customers? In practice, these two motives interact, making it hard to isolate the role each plays in generating price volatility. But doing so is important for understanding the magnitude of the micro-macro disconnect in price setting. If most of the micro volatility reflects PD, then it may not be relevant for understanding the aggregate dynamics of inflation. To make progress on this question, I assume that PD products feature policies that mostly consist of price *cuts* from a high modal price. Roughly one third of the OFP series and one half of MRP series have this property. But it turns out that the volatility of the data is not concentrated in the PD series. PD and non-PD series have similar policies, with two exceptions: PD policies last about twice as long, suggesting less fundamental volatility, and they have somewhat larger within-policy price changes, consistent with having large temporary discounts. Since the theory proposed does not include price discrimination, I only target the non-PD series.

The theory proposed to rationalize the empirical findings embeds costly information in an otherwise standard model of price setting. A continuum of heterogeneous firms set prices in the face of stochastic market conditions. All information about market conditions is available to these price-setters, but at a cost. Firms choose how much information to acquire, trading off pricing accuracy to save on information costs. Formally, each firm implements an optimal policy that specifies rules for acquiring information and for setting prices based on this information. Moreover, the policy itself can be revised, by paying a fixed cost. If it decides to review its policy, the firm pays a fixed cost which enables it to gather extensive—for simplicity, complete—information about the state of the economy and to reoptimize its policy. These reviews generate breaks in pricing, as seen in the data. In each period, the firm monitors its environment to decide (*i*) whether or not to pay the fixed cost to update its policy, and (*ii*) what price to charge in the period. These decisions are based on two imperfect signals, a review signal and a pricing signal. Both signals are modeled following the rational inattention literature (Sims, 2003), using entropy reduction as a measure of the informativeness of a signal (Shannon, 1948), and assuming that the cost of each signal is linear in this measure. The result is stochastic state dependence in both the review and the pricing decisions. How closely prices track the full-information profit-maximizing target price depends on firms' willingness to pay for more accurate signals, and on the frequency with which they choose to pay the (larger) fixed cost to learn the state and reset their policies.

The setup can be seen as capturing the interaction between headquarters (which chooses the policy) and the local branch (which sets prices day-to-day). Alternatively, it can be seen as a reduced-form representation of the relationship between the producer and the retailer: the overall policy is the result of (relatively infrequent) negotiations between the two parties, while the exact implementation (for instance, when to implement a sale) is up to the retailer.[3]

The theory delivers several novel results. First, it yields coarse prices in an infinite-horizon setting with Gaussian shocks. Even though the target price is continuously distributed, and its distribution is changing continuously, the firm reduces this complex state to a discrete, coarse approximation. The ability to occasionally undertake reviews is key for this result: The reviews prevent the distribution of target prices from becoming too dispersed. And having a manageable distribution of target prices to entertain between reviews means that the firm can afford to pick a pricing policy with a finite number of distinct price points. How many price points are charged with positive probability becomes a numerical question.

Second, the theory can generate heterogeneity in the complexity of pricing policies chosen by firms in different sectors or over time, consistent with the data. The model has a threshold cost of information that determines whether or not the firm acquires any pricing signals between reviews. If the sensitivity of profits to mispricing between reviews is low, relative to the cost of paying for the additional pricing signal, the firm chooses a single-price policy between reviews. It sets a price, and then it only monitors the evolution of market conditions to decide if it is time to change this price. Beyond this threshold, the cardinality of the pricing

[3] See Anderson, Jaimovich & Simester (2015) for the pricing practices of a U.S. retailer.

policy gradually increases, as does the accuracy with which the firm chooses which price to charge when. Hence, the coexistence of single-price, one-to-flex, and multiple-rigid-prices policies arises naturally if one allows firms to differ in the volatility of idiosyncratic shocks, in the costs of monitoring market conditions, or in the parameters governing the sensitivity of profits to mispricing. Quantitatively, the model matches the duration, coarseness, and volatility of the SPP and MRP policies documented in the empirical part of the paper.

Third, the review decision and pricing decision interact to determine how firms respond to shocks. As a result, matching the micro facts on pricing policies is essential for getting the aggregate dynamics right. In the general equilibrium parameterized to match the characteristics of SPP and MRP policies in the data, the theory predicts that multi-price firms are more responsive to an aggregate shock compared with single-price firms, especially on impact. This is consistent with the inflation dynamics observed in the data during the Great Recession. But is the difference purely reflecting the higher frequency of policy changes of MRP firms? And is the within-policy transitory volatility irrelevant to these firms' aggregate response? I find that the answer to both of these questions is no. The MRP firms respond differently to the shock not only because they update their policies more frequently, but also because they adjust prices between reviews. Filtering out the within-policy price volatility would overstate the degree of rigidity for the MRP series in the two-to-three months immediately following the shock (before many of the MRP firms have revised their policies), while under-stating it at longer horizons (by underestimating the degree of mistakes in within-policy pricing). Hence, the existence of transitory price volatility changes the impact and persistence of inflation's response to shocks.

Fourth, the severity of the information friction determines the degree of over-pricing relative to the full information benchmark. The profit function is asymmetric, generating larger losses from under-pricing (and having to meet the large resulting demand at high cost) than from over-pricing (and facing limited demand). As a result, information-constrained firms err on the side of over-pricing. Higher uncertainty makes the information problem more severe, generating even more over-pricing. Quantitatively, I estimate over-pricing of between two and five percentage points.

Finally, higher uncertainty also dampens firms' responses to shocks. This result stands in contrast to the predictions of full-information state-dependent pricing models, and it implies more effective monetary policy during high volatility periods. Intuitively, the result reflects the endogenous response of the firm's information acquisition strategy: When volatility rises, the firm increases its information expenditure, but it nevertheless faces higher posterior uncertainty. This effect may help explain recent inflation dynamics. The Great Recession was marked by both low aggregate demand and high volatility. These forces push prices in opposite direction: low aggregate demand induces the firm to reduce its prices, while higher volatility requires setting higher prices. This tension may help rationalize why inflation did not fall more during the Great Recession.

The empirical analysis contributes to a large literature on product-level price patterns (see Klenow & Malin, 2010 for a review).[4] This work has focused attention on transitory sales versus regular prices. I depart from that approach by interpreting both the transitory and the regular price levels as chosen to be jointly optimal, as part of an integrated policy. The resulting evidence of coarse policies is consistent with the simple price plans postulated by Eichenbaum, Jaimovich & Rebelo (2011) and generated here endogenously. Relative to this work, I also present evidence on patterns during the Great Recession.

The theory brings together different features from the costly information literature, primarily Reis (2006), Woodford (2009), and Matějka (2016), combining both lumpy and flow information acquisition, modeling a richer signal structure, and embedding the friction in a general equilibrium economy.[5] Integrating these features is important for reconciling high product-level pricing volatility with aggregate sluggishness. Quantitatively, I build on prior work by targeting a rich set of micro facts and aggregate dynamics. Lastly, I expand the discreteness results of Matějka (2016) beyond the static model with uniform shocks to a dynamic, infinite-horizon model with persistent Gaussian shocks. This shows that the rational inattention framework can generate discrete outcomes in a wide range of environments, which is promising for future work on lumpy adjustment in macroeconomics.

2 Empirical Evidence

In this section I discuss two sets of empirical results. First, I characterize the types of pricing policies observed over the entire sample period. Second, I focus on how these policies behaved during the Great Recession and the subsequent recovery, and what they implied for the dynamics of aggregate inflation.

2.1 Pricing Policies in Micro Data

The Data I use the *Retail Scanner Database* from The Nielsen Company (US), LLC. This database has weekly point-of-sale data on prices and quantities for products sold in stores from 90 retail chains across the United States. Product coverage represents about 27% of the total goods consumption measured by the *Consumer Expenditure Survey* of the Bureau of Labor Statistics (BLS). Categories include health and beauty care, food, beverages,

[4] See also Bils & Klenow (2004), Klenow & Willis (2007), Klenow & Kryvtsov (2008).

[5] Other models of price setting with endogenous information acquisition include Maćkowiak & Wiederholt (2009, 2015), Paciello (2012), Paciello & Wiederholt (2014) and Pasten & Schoenle (2016).

general merchandise, and others.[6] I limit the sample to the store with the largest number of observations from each chain.[7] Some series have missing observations. I keep only series with at least 52 contiguous observations. The resulting sample has weekly observations on more than one million unique store-UPC pairs, from January 2006 through December 2015.

The advantages of these data are the high frequency of observations, the relatively long time series for individual products, and the large number of products within the categories and across locations. Conversely, the micro data underlying the BLS's *Consumer Price Index* (CPI) has monthly or bimonthly sampling, high product turnover rates, and narrower sampling within product groups and across regions. The drawback of the Nielsen data is the narrow coverage of product categories. Nevertheless, it covers products whose prices are highly volatile and exhibit precisely the sharp, transitory price swings that have been at the forefront of the recent price dynamics literature. The median weekly frequency of price changes is 24.6% and the median absolute size of price changes is 11.9%. For comparison, the monthly frequency and the size of price changes for products underlying the CPI average 10.6% and 9.6% over the 2006-2014 period (Nakamura, Steinsson, Sun & Villar, 2018).[8]

The Break Test The empirical method identifies pricing plans at the product level by looking for breaks in individual price series. To identify the break points, I adapt the Kolmogorov-Smirnov test, which tests whether two samples are drawn from the same distribution. I interpret each break as the transition to a new plan, characterized by a new distribution of prices.

Building on work that estimates the location of a single break in a series (Deshayes & Picard, 1986, and Carlstein, 1988), I modify the method to identify an unknown number of breaks at unknown locations in a series. I use an iterative procedure similar to that of Bai & Perron (1998), who sequentially estimate multiple breaks in a linear regression model. I first test the null hypothesis of no break in a series; upon rejection, I estimate the location of the break; I then iterate on the two resulting sub-series until I fail to reject the null of no

[6] Data provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business, http://research.chicagobooth.edu/nielsen. The conclusions drawn from the Nielsen data are my own and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. The data have also been used by Beraja, Hurst & Ospina (2018) to analyze dynamics in regional price indices.

[7] DellaVigna & Gentzkow (2017) document near-uniform pricing within chains, so I use one store per chain. As is common in the literature, I exclude the Deli, Packaged Meat, and Fresh Produce departments.

[8] I exclude price changes less than 1% in absolute value (11% of all changes). In the full sample, the median frequency and size of price changes are 28% and 11%. However, as argued by Eichenbaum, Jaimovich, Rebelo & Smith (2014), very small price changes may reflect measurement error, since in these data, a price observation is the volume-weighted average price of the product in each week. Prices reflect bundling (e.g. 2-for-1 deals) and discounts associated with the retailer's coupons or loyalty cards. Variation in bundling or in the use of such discounts across weeks may induce spurious small price changes.

break. The critical value used to reject the null of no break is determined via simulations, starting from the asymptotic critical values provided by Deshayes & Picard (1986), which are valid for the test of a single break on i.i.d. data drawn from continuous distributions.[9]

The test's usefulness depends on its ability to correctly identify the timing of breaks. I find that the break test correctly identifies breaks 91% of the time in simulated data. It finds the exact location of the break 94% of the time and is off by two periods in the remaining cases. In simulated series restricted to have at least five observations between breaks, the test finds virtually all breaks. It loses power for policies lasting less than five weeks, because there are not enough data points to be confident about the distribution generating them.

Applied to the Nielsen product-level series, this procedure identifies interesting patterns of across-policy and within-policy volatility. Table I reports the key facts.

Stickiness The first empirical result is that the identified pricing policies change infrequently. Breaks in the price series typically occur every 7.7 months, and most policies last at least 4.5 months, even though raw prices change every three-to-four weeks. For comparison, papers that seek to filter out transitory price volatility report the duration of regular or reference prices ranging from 7.8 months to 12.7 months in grocery store data, and from 6.7 months to 14 months in the CPI.[10] This variation across studies even when using similar data highlights the fact that measures of stickiness are sensitive to the definition of permanent versus transitory price changes and to the filters implemented to identify them.[11] An advantage of the break test over the filters is precisely the fact that it sidesteps the need to take a stand on how to define and identify regular versus transitory price changes, which is the source of a big portion of the dispersion in estimates in the existing literature.

Volatility Between consecutive breaks, the prices charged are quite volatile. The median weekly frequency of within-policy price changes is 24.6%, consistent with existing work that has identified frequent transitory price volatility accompanying the slower dynamics of regular or reference prices. The data also feature large price changes both within and across

[9] I simulate data as a mixture of processes that represent commonly observed pricing patterns: sticky prices, sticky prices with transitory deviations of variable sign, size, and duration, and sticky plans with a variable number of prices. The simulation targets the range of frequency and size of price changes observed in the micro data. The critical value is determined by trading off power against false positives in simulated data. The online appendix details the method and its performance across the different processes.

[10] Midrigan (2011), Kehoe & Midrigan (2015) and Eichenbaum et al. (2011) report statistics for grocery store data and Klenow & Kryvtsov (2008), Nakamura & Steinsson (2008), and Kehoe & Midrigan (2015) for CPI data, using different filters. I report the monthly implied duration = -1/ln(1-median monthly frequency) for all studies, to limit bias due to the censoring of individual price series.

[11] In particular, v-shaped filters tend to yield significantly lower duration estimates, because they only allow for transitory price *decreases* from a rigid mode, whereas I find that transitory price increases from the rigid mode occur in more a third of the policy realizations in my sample.

Table I: Characteristics of Pricing Policies

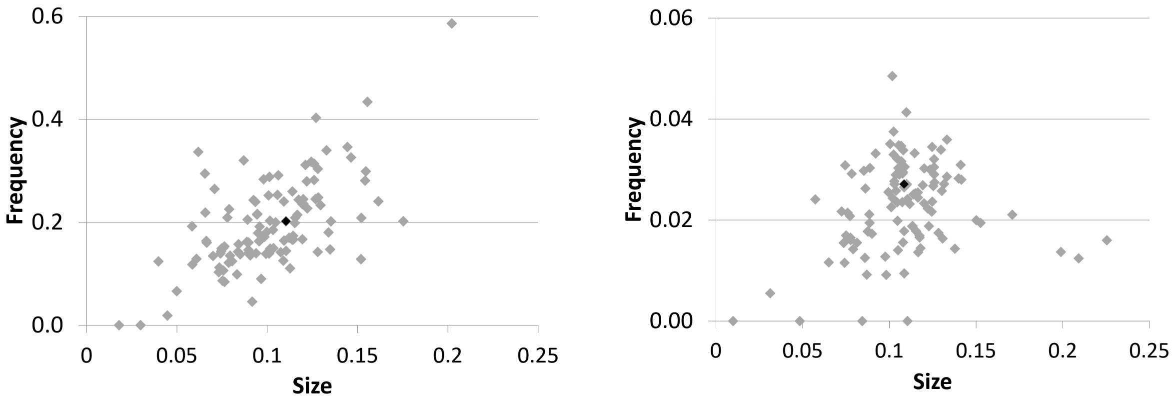
	All	Single-price	One-to-flex	Multi-rigid
Fraction of series (%)	100	12.0	28.5	59.5
Monthly frequency of policy changes (%)	12.2	7.8	17.3	12.0
Implied policy duration (months)	7.7	12.3	5.3	7.9
Freq. of weekly price changes within (%)	24.6	0.6	15.0	35.8
Size of price changes within (%)	11.9	5.8	9.9	13.6
Size of shift across (%)	11.3	8.5	11.7	11.5
Policy cardinality	3	1	3	4

Note: Nielsen Retail Scanner data, 2006-2015. *Implied policy duration* is the duration implied by the median monthly frequency of policy changes. *Frequency of weekly price changes within* is the median weekly frequency with which prices change between breaks. *Size of price changes within* is the absolute value and is non-zero for SPP because the category allows for rare deviations from the modal price. *Size of shift across* is the median absolute change in the weighted average price across policy realizations.

policy changes. The median absolute size of within-policy price changes is 11.9% and the median shift in prices across consecutive policy realizations is 11.3%.[12] These magnitudes are consistent with prior evidence that prices often change by amounts that are much larger than what is needed to keep up with aggregate inflation. Instead, they point to the importance of idiosyncratic drivers of price adjustment (e.g. Golosov & Lucas Jr, 2007, and Klenow & Kryvtsov, 2008). The novelty here is the distinction between within-policy price changes and shifts in the price levels across policies. This distinction is useful, since it can identify the role of different drivers of price variability. Within-policy volatility may be primarily driven by transitory shocks or price discrimination motives, while the shift in prices across policies may be driven by more persistent shocks.

The patterns of volatility can also identify different frictions and sources of heterogeneity in price adjustment. Figure 1a shows a scatter plot the frequency and size of within-policy price changes, and Figure 1b shows a scatter plot of the frequency of policy changes versus the size of shifts across policies, for the different product groups in the sample. The positive correlation between the frequency and the size of adjustment in these panels is difficult to reconcile with theories of price rigidity driven by heterogeneous menu costs, which would be

[12] The policy shift is obtained by computing the average weighted price within each policy realization, and then computing the absolute value of the change in this average across consecutive policies.



(a) With-policy price changes

(b) Across-policy shifts

Figure 1: Frequency and size of within and across policy price changes across product groups

Note: Nielsen Retail Scanner data, 2006-2015. Panel (a) plots the frequency against the absolute size of within-policy price changes. Panel (b) plots the frequency of policy changes against the absolute size of the shift in the average price across consecutive policies. Points indicate means at the product group level. The expenditure-weighted medians for the full sample are in black.

generate a negative correlation between the size and the frequency of adjustment. Instead, these plots suggest heterogeneity in the volatility of market conditions that firms face. Some products rarely update their prices, and even when they do, they change by modest amounts, while others change prices quite frequently and also by large amounts.

Coarseness Although they last a fairly long time and display volatile prices, policy realizations also exhibit coarse pricing. The median number of distinct prices per policy realization is three, and the large majority of policies have less than six price points.[13] This finding points to the "disproportionate importance" of a few price points at the policy level, consistent with similar evidence at the series level documented by Klenow & Malin (2010) using the micro data underlying the CPI. This coarseness is also what helps identify break points. What changes systematically across policy realizations is the support of the price distribution; there is no consistent change in the shape or the cardinality of the distribution.

Overall, firms appear to have flexibility in the *timing* of price adjustment, but rigidity in the *level* to which the price adjusts. This combination is at odds with virtually all models of rigid pricing, in which, conditional on deciding to adjust, firms flexibly choose a new price, thus ruling out rigid price levels (insofar as market conditions evolve smoothly). It is also at odds with models in which firms choose deterministic price paths that generate continuous gradual adjustments. The theory proposed in the next section generates the coexistence of

[13] These statistics are based on average weekly prices, so they likely understate coarseness.

these two features of the data.

Policy Heterogeneity Prior work has documented substantial heterogeneity in the frequency of price changes across goods (e.g., Nakamura & Steinsson, 2008). I find that more generally, there is heterogeneity in the types of pricing policies that products exhibit. Moreover, the different policy types can be matched to some popular models of price setting, while ruling out others.

I classify the policy realizations between consecutive breakpoints in terms of the rigidity in the observed price levels. I then classify each product series in terms of the types of policy realizations observed over the life of the series.[14] Figure 2 shows the incidence of policy types across product groups, and Table I presents key statistics.[15]

Single-Price Policies The workhorse Calvo or menu cost models of rigid price setting generate sequences of *single-price policies* (SPP). Each policy realization consists of a single price, and a break is a shift to a new price. In the data, only about 2% of the series are characterized by such clean single-price plans. Hence, I relax the definition of SPP series to allow for occasional deviations from such rigidity, recognizing that such infrequent deviations are likely not a systematic feature of the firm's pricing policy and may reflect some degree of measurement error. Specifically, I identify a realization of a single-price policy features as a single sticky price with at most one deviation between two consecutive breaks, and I categorize as effectively single-price products all products for which at least 90% of observations fall inside such policy realizations. Approximately 12.0% of products are SPP products defined in this way.

The prices of SPP goods adjust much less frequently, and by less when they do adjust: the median policy duration is 12.3 months versus 7.7 months for all products, and the median shift in prices is 8.5% versus 11.3% for all products. These goods appear to face a relatively low volatility of their desired price that does not warrant designing complex pricing policies or undertaking large or frequent price changes.

One-to-Flex Policies Motivated by the high incidence of transitory price changes in the data, more recent pricing models (Kehoe & Midrigan, 2015 and Guimaraes & Sheedy, 2011) feature a rigid *regular* or *reference* price accompanied by transitory deviations to and from it,

[14]For this classification, I assume that the type of policy employed for a particular product does not change over the sample period, and I test this assumption in Section 2.2.

[15]The online appendix reports statistics at the policy-product level, which are consistent with those at the series level. It also reports statistics based on an alternative series classification method; alternative critical values for the break test; and alternative identification of breakpoints, using the rolling mode filter of Kehoe & Midrigan (2015).

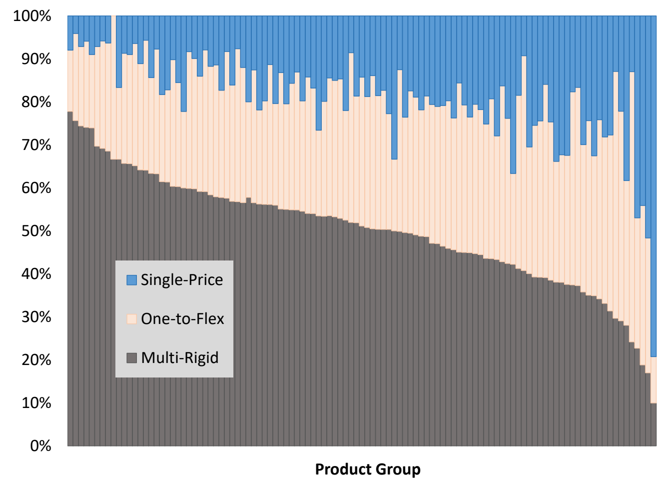


Figure 2: Classification of series by type of pricing policy, across product groups

Note: Nielsen Retail Scanner Data. Percent of series of each type in each product group.

in a *one-to-flex* pattern.[16] I identify the OFP series in the data as series for which a plurality of policy realizations feature the OFP pattern.[17] In the data, 28.5% of products are OFP series. The median policy duration is much shorter, at 5.3 months, and the median shift in average prices across policy realizations is 11.7%. However, the policies themselves are not very volatile or complex. They display two or three distinct prices, and the median frequency with which prices adjust inside policies is only 15.0% (versus 24.6% for all products). The relatively high across-policy volatility together with the lower within-policy volatility suggest that these products face a higher volatility in their desired price than the SPP goods, but also a high cost of implementing complex policies, which induces them to instead update their policies more frequently.

Multi-Rigid Policies Underscoring the rigidity in price levels beyond that of the modal price of each price plan, 59.5% of products are characterized by policies with multiple rigid prices. These are series for which a plurality of policy realizations feature at least two prices

[16]Kehoe & Midrigan (2015) assume that firms can "rent" a one-period price deviation for free, but must pay a cost to change the price permanently. Alternatively, Guimaraes & Sheedy (2011) develop a price discrimination model that features a two-price distribution in the steady state; with shocks, they assume that the high price changes a la Calvo, while the low price changes freely.

[17]Since price data are noisy, I allow for some flexibility in categorizing products, by allowing for SPP or MRP policy realizations inside OFP series. In the online appendix, I report results for an alternative classification in which OFP series do not exhibit any MRP realizations. This reduces the incidence of products categorized as OFP and makes these products look very similar to SPP goods.

that are revisited over the life of the policy.[18] The median policy duration for these products is 7.9 months, but only three to four distinct prices are typically charged over the life of the policy. These products exhibit high volatility: The median shift in prices across policy realizations is 11.5%, the median absolute size of within-policy price changes is 13.6%, and the median frequency of within-policy price changes is 35.8%. These statistics suggest that these products face highly volatile market conditions, and they adjust by choosing more complex — though nevertheless coarse — pricing policies.

The prevalence of MRP goods in the data poses a challenge to existing theories of price rigidity. It is instead consistent with the hypothesis that firms set a small menu of prices which they update relatively infrequently. The theory developed in Section 3 uses costly information to generate such plans endogenously.

Price Discrimination Policies Overall, series overwhelmingly feature price changes between policy shifts. What drives this within-policy volatility? The obvious reasons are responding to shocks and attempting to price discriminate among heterogeneous customers. In practice, these motives interact, making it difficult to isolate how important each one is for price volatility. But disentangling these factors is important for quantifying the severity of the disconnect between product-level price volatility and aggregate sluggishness in inflation. If price discrimination is a dominant factor, then the product-level price volatility may be less relevant for the aggregate dynamics of inflation. I estimate the incidence of *price discrimination policies* (PDP) by defining them as policies in which the maximum price is also the mode. Among one-to-flex and multi-rigid series, I label a series as price discriminating if a majority of its policy realizations are PDP. This definition is consistent with models of price discrimination that feature a mass point at the high price of the pricing policy (e.g., Guimaraes & Sheedy, 2011). Roughly one third of the OFP series and one half of MRP series fit this description. Table II reports the statistics for the price discrimination series separately from the non-price discrimination series. PD series feature much longer policy durations (suggesting lower fundamental volatility) and somewhat larger within-policy price changes (consistent with having large discounts to attract bargain hunters). But the remaining non-PD series are also highly volatile, exhibiting frequent and large within-policy price changes. I conclude that the data continue to point to a large micro-macro volatility gap, which requires a model of price setting that divorces product-level volatility from aggregate price flexibility. The information friction modeled in the next section closes this gap by generating noisy pricing that tracks market conditions imperfectly.

[18] A price level is "revisited" if the price returns to that level before a break occurs in the series.

Table II: Price Discrimination Policies

	One-to-flex		Multi-rigid	
	PD	Non-PD	PD	Non-PD
Fraction of all series (%)	10.2	18.3	29.5	30.0
Monthly frequency of policy changes (%)	11.6	21.3	9.1	14.8
Implied policy duration (months)	8.1	4.2	10.4	6.2
Freq. of weekly price changes within (%)	14.9	15.0	33.3	38.6
Size of price changes within (%)	11.8	9.1	15.3	12.2
Size of policy shift (%)	11.9	12.6	12.4	10.4
Policy cardinality	3	2	4	4

Note: Nielsen Retail Scanner data. Statistics for one-to-flex and multi-rigid series that are price discrimination (*PD*) and non-price discrimination (*Non-PD*) series, where PD series are defined as series in which a majority of policy realizations have the maximum price equal to the modal price.

Break Test versus Filters How much do we gain by allowing for non-parametric changes in the distribution of prices charged? Conceptually, the break test is more flexible in its identification of breaks than filters that identify changes in a particular statistic (such as the modal or the maximum price charged). This flexibility allows me to first identify breaks in price series, and then investigate what aspects of the distribution change across breaks. Simulations suggest that the break test is preferable: while each filter does particularly well on specific data generating processes, the break test does well across different processes, especially when the processes are characterized by random variation in the duration of both regular and transitory prices.[19] By using information about the entire distribution of prices, the break test also has more accuracy in detecting the *timing* of breaks compared with methods that focus on a single statistic. While the existing literature has focused more on the median duration of regular prices, accurately identifying the timing of breaks is particularly important for characterizing within-policy volatility and the responsiveness to shocks. Statistics such as the number of distinct prices charged, the prevalence of the highest price as the most frequently charged price, or the existence of time-trends between breaks are also sensitive to the estimated location of breaks.

[19]The online appendix compares the performance of the break test to that of three filters in simulated as well as actual data. The rolling model filter proposed by Kehoe & Midrigan (2015) gives results that are closest to the break test in terms of both accuracy in simulated data and synchronization in actual data.

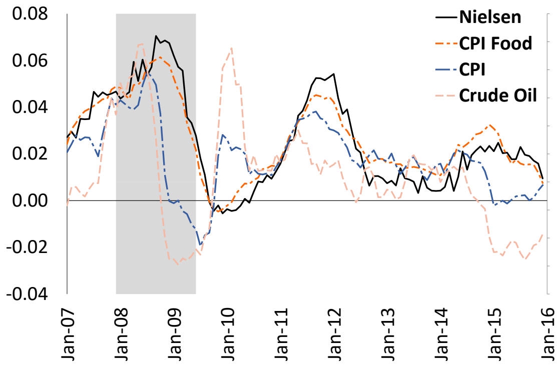
The break test isolates the more persistent changes in pricing patterns from the transitory pricing dynamics. But it itself does not introduce artificial rigidity in the measurement of policy durations, which could be a concern. In simulations in which prices change flexibly every three-to-four weeks, the test would conclude that policies last on average six weeks. This duration is much lower than what I find in the data, where the large majority of policies last at least 20 weeks. Hence, the method's low power for very short-lived policies does not appear to be a constraint for finding break points in the actual data.

2.2 Dynamics During the Great Recession

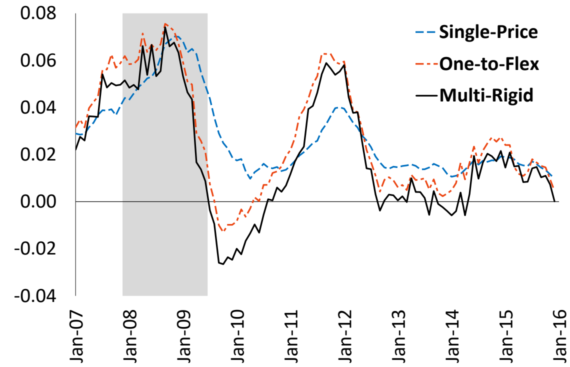
Making the distinction between different types of policies and between policy changes and raw price changes is useful for disentangling the dynamics of inflation during the Great Recession and its aftermath.

Inflation Figure 3a shows the aggregate dynamics, plotting the annualized monthly inflation rate of the Nielsen sample compared with that of the CPI, and with crude oil price inflation. The Nielsen inflation rate closely tracks the Food and Beverages CPI inflation rate (the correlation between the two series is 96%; see also Beraja et al., 2018). Interestingly, these two series started diverging significantly from the overall CPI inflation in October 2008, at the height of the recession. They fell much more slowly and continued to diverge over the sample period, with the exception of a period of more stable oil prices, between early 2012 and early 2014. Much of the gap between these series and the overall CPI inflation rate reflects the fact that CPI inflation has tracked crude oil price inflation much more closely than the Nielsen inflation rate (79% versus only 3% correlation). This is surprising since history would predict Nielsen to be more correlated with oil, not the overall CPI basket. But in this decade, Nielsen prices (and Food CPI more generally) were more rigid during the recession and its immediate aftermath, and also more inflationary starting in 2011. Had oil prices been less volatile, the missing disinflation puzzle of the Great Recession might have been even more severe in the aggregate data.

Figure 3b decomposes the Nielsen inflation rate into the inflation rates for the three types of products — single-price, one-to-flex and multi-rigid. There are stark differences in the degree of state-dependence across the different policy types. All three inflation rates moved largely in tandem at the beginning and the end of the sample, suggesting limited divergence in "tranquil" times. But they diverged significantly in the crisis and its immediate aftermath. During this period, the inflation rate for MRP products fell much more than for SPP products. In fact, single-price products continued to raise prices throughout, while MRP goods cut prices, and MRP inflation fell to a low of −2.7%. Hence, MRP goods, which



(a) Nielsen and Aggregate



(b) Nielsen, by policy type

Figure 3: Annualized inflation in Nielsen and the CPI versus crude oil

Note: Nielsen Retail Scanner and Bureau of Labor Statistics data. Crude oil inflation is re-scaled for comparability. The shading marks the Great Recession.

likely face more volatile market conditions in general, also responded more aggressively to the aggregate shocks during this volatile period.

Table III assesses these differences more formally, exploiting cross-sectional variation at the national and state levels. First, I compute the monthly inflation rate at the module level,[20] for products of each policy type across all locations, and regress it on monthly national unemployment, using the specification

$$\pi_{ikt} = \sum_{h=1,2,3} (\alpha_h + \beta_h U_t) D_h + \delta_i + \gamma_t + \lambda_{it} + \epsilon_{ikt}, \tag{1}$$

where π_{ikt} is the inflation rate across products of policy type k in module i and month t, U_t is the unemployment rate, D_k is a policy type dummy, δ_i, γ_t and λ_{it} are module, month, and module-by-month fixed effects. The table reports results with and without the time fixed effects. The sensitivity of inflation to unemployment is significantly higher for MRP series than it is for SPP series, and it is not driven by variation in specific modules over time. However, the national data are exploiting essentially a single episode of high unemployment. So I also report results using state-level inflation and state-level unemployment. I expand the sample to include multiple chains in each state, and I keep data from the largest store within each chain and state. This increases the sample size to more than nine million observations, and offers more variation in both inflation and unemployment rates. Using state-level monthly unemployment as a measure of local demand conditions, I regress monthly inflation π_{sikt} — now defined at the state-by-module-by-policy level — on local demand, also adding state,

[20]Nielsen groups products into roughly 1,000 product modules.

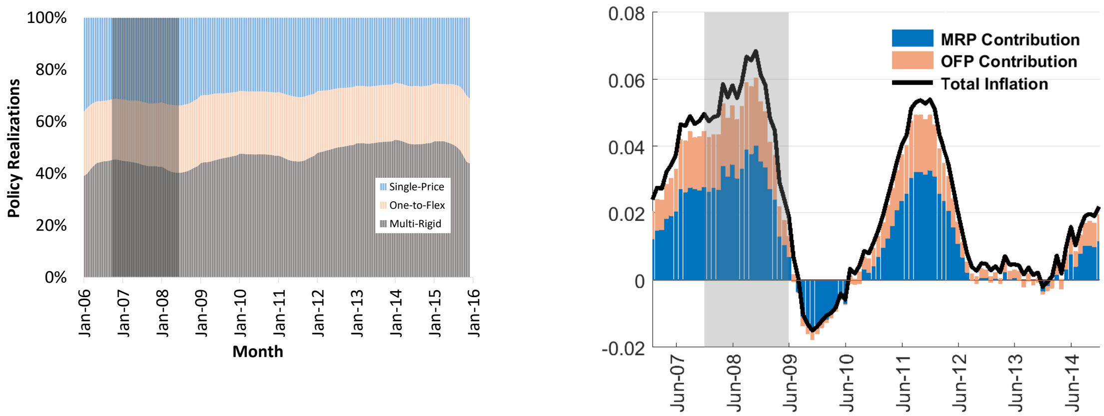
Table III: Sensitivity of Inflation to Unemployment by Policy Type

	National		State-level	
	(1)	(2)	(3)	(4)
Unemployment	-0.046	–	0.015	–
	(0.038)	(.)	(0.041)	(.)
Unemployment x OFP	-0.042	-0.089	-0.071^{**}	-0.075^{***}
	(0.050)	(0.049)	(0.027)	(0.027)
Unemployment x MRP	-0.098^{*}	-0.146^{**}	-0.094^{***}	-0.099^{***}
	(0.047)	(0.045)	(0.026)	(0.027)
Module FE	Yes	-	-	-
Module x Month FE	-	Yes	Yes	-
State FE	-	-	Yes	-
State x Module x Month FE	-	-	-	Yes
Observations	254,600	244,820	9,626,039	9,079,654
R^2	0.0608	0.5304	0.2148	0.5669

Note: Nielsen Retail Scanner Data. The first two columns report regressions of module-policy inflation on national unemployment. Standard errors are in parentheses, clustered at the module level. Two-way clustering by module and month (omitted) reduces significance from the 1% to the 5% level for MRP series. The last two columns report regressions of module-policy inflation by state on state-level unemployment as a proxy for local demand. Standard errors are in parentheses, two-way clustered at the state and month level. All regressions also include policy dummies and fixed effects, as indicated (not reported). $^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05$.

module, and time fixed effects. The specification sweeps out the common variation coming from the Great Recession, but nevertheless, MRP inflation responds significantly to local unemployment, while SPP inflation has no meaningful response.

The heterogeneous responsiveness to the state of the economy documented here supports the results of Gilchrist, Schoenle, Sim & Zakrajšek (2017), who find that at the peak of the crisis, firms operating in competitive markets lowered their prices significantly, relative to firms operating in less competitive markets. The information-based theory presented in the next section predicts this connection: firms that operate in more volatile or more competitive markets choose more complex pricing policies and respond to shocks more aggressively, while



(a) Policy realizations over time

(b) Contributions of policy types to inflation

Figure 4: Policy choices and their contribution to inflation over time

Note: Nielsen Retail Scanner and Bureau of Labor Statistics data. Panel (a) plots the type of policy realizations over time. Panel (b)plots the contribution of MRP and OFP series to total inflation. The difference between total inflation (marked by the black line) and the sum of MRP and OFP contributions is the contribution of SPP series.

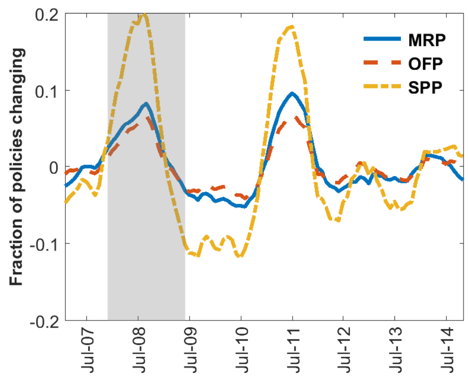
firms that choose simpler policies also adjust more sluggishly to the state of the economy.

Importantly, these findings underscore the value of studying price data in its entirety, without eliminating transitory price volatility. Transitory volatility is crucial to pinning down the type of pricing policy of different products and, in turn, the type of policy is correlated with how these firms respond to shocks, thereby affecting aggregate inflation dynamics. Splitting the data by frequency of policy reviews — rather than by type of policy — would generate a much less significant relationship between inflation and unemployment across all groups, because it would mix the longer duration single-price series with the longer duration multi-rigid series, which in fact have different cyclical properties.

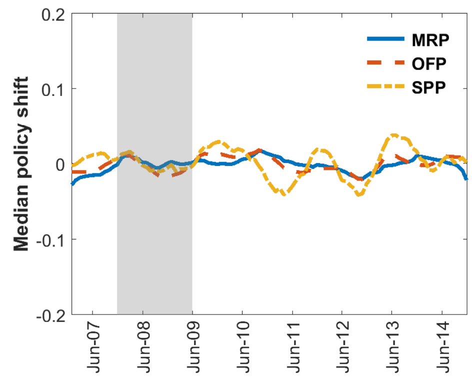
Cyclical Policy Choice? So far, we have seen that MRP products reacted more aggressively to the state of the economy during the recession, while SPP products barely responded. In establishing this result I have assumed that products do not change their policy type over time, so that they can be assigned once and for all to a particular category. However, both the type and the statistics of the policies being realized may vary over time.

How important are changes in the types of policies being realized over time? As shown in Figure 4a, there is some variation in the incidence of different types of policies, with MRP realizations increasing slightly, at the expense of single-price policies. This trend supports the notion that pricing has become more complex in the U.S. in recent decades.[21] But the

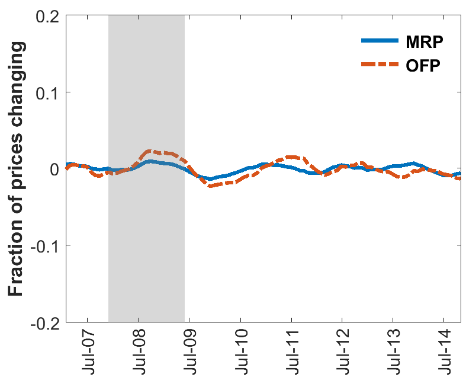
[21] E.g., Nakamura et al. (2018) document an increase in the incidence of temporary sales over time.



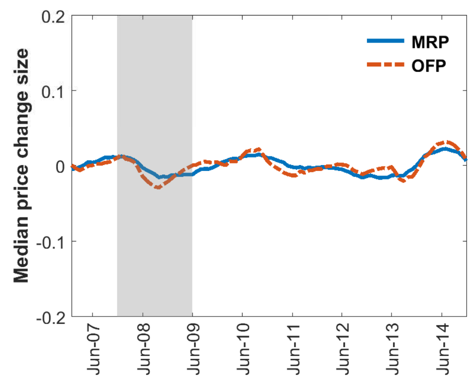
(a) Frequency of policy changes



(b) Size of policy shifts



(c) Frequency of price changes



(d) Size of price changes

Figure 5: Policy and price adjustment during the Great Recession

Note: Nielsen Retail Scanner Data. The top panels show time series for the deviations from the trend in the fraction of policy changes over time (Panel a) and in the median absolute shift in prices across policies (Panel b), by policy type. The shift is obtained by computing the change across policies in the weighted average price. The bottom panels show time series for the frequency of within-policy price changes (Panel c) and for the size of within-policy price changes (Panel d), by policy type. These are seasonally adjusted weekly figures, averaged to monthly values and filtered with a Baxter-King bandpass filter with parameters 12, 96, 18.

increase is not monotonic. Surprisingly, the Great Recession saw a *decrease* in multi-rigid policy realizations and an increase in single-price policies. The same pattern occurred in 2011, another period of heightened volatility. Over this decade, it seems that firms' policies have become more complex, except in periods of uncertainty, during which firms seem to favor implementing simpler, single-price policies.

Nevertheless, the variation in the types of policies realized over time has only a modest effect on inflation dynamics. The main driver of inflation is the volatility of the MRP series. Figure 4b shows the contribution of SPP, OFP, and MRP series to aggregate inflation. The MRP series are the most important not only because they have a large share in the overall number of series, but also because they feature the most volatile inflation. By contrast, SPP series have only a marginal contribution to aggregate inflation dynamics.

Cyclical Policy Adjustment? How do the prices within the different policy types adjust to generate the inflation dynamics seen in the aggregate? For each policy type, I decompose inflation dynamics into the contribution coming from policy adjustments (frequency size of shifts across policies) and from within-policy price changes. During the recession, adjustments overwhelmingly reflect across-policy rather than within-policy changes. As shown in Figure 5, the key adjustment margin is the frequency of policy changes, which rose substantially for all product types. This supports the hypothesis of at least partial state-dependence in policy adjustment. The rate of policy changes increased particularly sharply for single-price products, with a 20 percent increase at the height of the recession. But this increase did not translate into much flexibility in the price index for these firms, which, as we have seen, had a muted response to the Great Recession.

The bottom panels of Figure 5 document the patterns over time for within-policy volatility. The percent changes in the within-policy rates of price adjustment are about an order of magnitude smaller than changes in the rate of policy adjustments. Likewise, the absolute size of within-policy price changes did not change significantly, decreasing by less than five percent. A possible explanation for these patterns is that the heightened uncertainty associated with the Great Recession led firms to keep revising their pricing plans instead of making them more complex. This interpretation is bolstered by the increase in the rate of policy adjustments that took place in 2011, which was another period of increased uncertainty due to the Euro zone crisis, the U.S. fiscal policy crisis, and rising and highly volatile oil prices.[22]

3 Theory

The empirical evidence supports a theory of price setting that generates coarse, infrequently updated price plans. In this section, I develop a theory of information acquisition that can generate such price plans endogenously.

3.1 The Agents

The economy consists of a fully informed representative household, a continuum of information-constrained producers, and a government that follows an exogenous policy.

[22]This evidence is also consistent with that of Anderson, Malin, Nakamura, Steinsson & Simester (2017), who find that an increase in oil prices in the 2007-2009 period had a significant effect on the frequency of *regular* prices posted by a particular retailer. Berger & Vavra (2018) and Nakamura et al. (2018) document countercyclicality in the frequency of regular price changes in the CPI in recent decades; here I emphasize the role of volatility even in the absence of a recession.

Households The household's problem is standard. The household has full information and chooses paths for consumption, labor supply, money, and bonds to solve

$$\max_{\{C_t, C_{it}, H_{it}, M_t, B_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{1}{1+\nu} \int_0^1 H_{it}^{1+\nu} di \right] \tag{2a}$$

$$M_t + B_t \leq M_{t-1} + (1 + i_{t-1}) B_{t-1} + \int_0^1 W_{it} H_{it} di + \int_0^1 \Pi_{it} di + T_t - P_{t-1} C_{t-1}, \tag{2b}$$

$$P_t C_t \leq M_t, \tag{2c}$$

$$C_t \equiv \left[\int_0^1 [A_{it} C_{it}]^{(\varepsilon-1)/\varepsilon} di \right]^{\varepsilon/(\varepsilon-1)}, \tag{2d}$$

where A_{it} is a good-specific preference shock, H_{it} is the differentiated labor supplied to each firm i, W_{it} is the nominal hourly wage of firm i, Π_{it} is the dividend received from firm i, T_t is the net monetary transfer received from the government, B_t is the amount of risk-free nominal bonds held in the period, i_t is the risk-free nominal interest rate on these bonds, M_t is money holdings, $\beta \in (0, 1)$ is the discount factor, $\varepsilon > 1$ is the elasticity of substitution, $\sigma > 1$ is the constant relative risk aversion parameter, $\nu \geq 0$ is the inverse of the Frisch elasticity of labor supply, and $P_t \equiv \left[\int_0^1 (P_{it}/A_{it})^{1-\varepsilon} di \right]^{1/(1-\varepsilon)}$ is the aggregate price index. The optimality conditions are standard and shown in the online appendix.

Government For simplicity, the government follows an exogenous policy. The net monetary transfer in each period is equal to the change in money supply, $T_t = M_t^s - M_{t-1}^s$, where the log of money supply evolves according to $\log M_t^s = \log M_{t-1}^s + \eta_t$, $\eta_t \overset{i.i.d.}{\sim} h_\eta$.

Firms A continuum of monopolistically competitive firms produce differentiated goods using the production function $Y_{it} = H_{it}^{1/\gamma}/A_{it}$, where H_{it} is the differentiated labor input, A_{it} is the firm-specific inverse of productivity, and $\gamma \geq 1$ captures the returns to scale in production. The stochastic variable A_{it} represents the effort required to produce the good and also increases the utility from consuming it.[23] The law of motion for this quality term is $\log A_{it} = \log A_{i,t-1} + \xi_{it}$, with $\xi_{it} \overset{i.i.d.}{\sim} h_\xi$. Excluding information costs, nominal profit is $\Pi_{it} = P_{it} Y_{it} - W_{it} H_{it}$. The profit maximizing full information flexible price is $X_{it} \equiv A_{it} M_t / Y^*$, where Y^* is the associated equilibrium output, $Y^* \equiv [(\varepsilon - 1)/(\varepsilon\gamma(1 + \nu))]^{1/(\sigma+\gamma(1+\nu)-1)}$.[24]

[23] The assumption that this term enters both the household's demand and the firm's cost implies that the firm's profit is shifted in the same way by the aggregate nominal shock and by this idiosyncratic shock, which enables a reduction in the state space of the problem. See also Midrigan (2011) and Woodford (2009).

[24] The online appendix derives this and all subsequent results that are omitted here for brevity.

3.2 The Firms' Information Problem

Monitoring the state of the economy is costly for firms, but they can choose how much attention to pay to market conditions. Each firm chooses a policy that specifies a menu of prices and a rule that determines which price to charge in each period and state of the world. How many prices are on the menu and how sensitive the rule is to market conditions depend on the firm's willingness to acquire more information in order to make its prices track the full information target price more closely. Moreover, motivated by the evidence of breaks in product-level price series, I assume that firms can revise their policies, subject to a fixed cost. This means that in addition to deciding which price to charge in each period, firms must also decide whether or not to undertake a policy review. How much information about market conditions to acquire in order to make this decision is also their choice, depending on how valuable it is to have accurately-timed policy reviews.

Objective Each firm maximizes its discounted expected profits net of the monitoring and policy review costs. The fixed cost of policy reviews makes the firm's problem dynamic. Let π_{it} denote a firm's per-period profit in units of marginal utility, excluding information costs. Profit in the economy with costly information can be written as a function of the gap between a firm's actual price and the frictionless target X_{it}, and of the gap between actual output and the frictionless level of output Y^*:

$$\pi_{it} = (Y^*)^{1-\sigma} \left[\left(\frac{P_{it}}{X_{it}}\right)^{1-\varepsilon} \left(\frac{Y_t}{Y^*}\right)^{2-\varepsilon-\sigma} - \frac{\varepsilon - 1}{\varepsilon\gamma(1+\nu)} \left(\frac{P_{it}}{X_{it}}\right)^{-\varepsilon\gamma(1+\nu)} \left(\frac{Y_t}{Y^*}\right)^{\gamma(1+\nu)(1-\varepsilon)} \right], \quad (3)$$

where aggregate output relative to frictionless output depends only on the joint distribution of prices and targets in each period, after firms have made all their decisions:

$$Y_t = Y^* \left[\int_0^1 (P_{it}/X_{it})^{1-\varepsilon} \, di \right]^{-1/(1-\varepsilon)}. \quad (4)$$

The information-constrained firm chooses a pricing and reviewing policy that solves

$$\max_{\left\{P_{it}, I^p_{it}, I^r_{it}, \delta^r_{it}\right\}} E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_{it} - \theta^p I^p_{it} - \theta^r I^r_{it} - \kappa \delta^r_{it}\right], \quad (5)$$

where I^p_{it} is the quantity of information acquired in period t in order to make the pricing decision, with unit cost $\theta^p > 0$, I^r_{it} is the quantity of information acquired to decide whether or not to review the policy, at a unit cost $\theta^r > 0$, δ^r_{it} is equal to 1 if the firm reviews its policy in period t and 0 otherwise, and $\kappa > 0$ is the fixed cost associated with a policy review. Payment of this fixed cost enables the firm to obtain complete information about

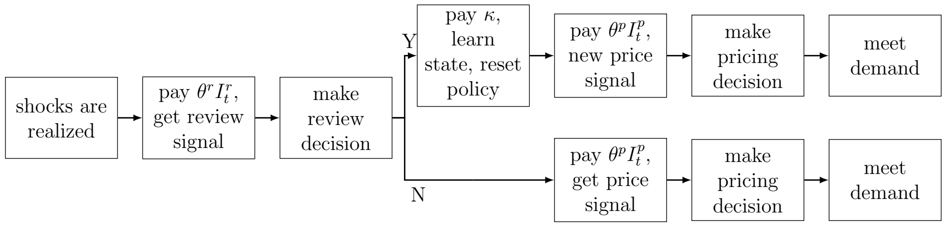


Figure 6: Sequence of events in each period of the model.

the economy at the time of the review, as in Reis (2006) and Woodford (2009).[25]

Monitoring Market Conditions Figure 6 presents the timeline in each period. The firm monitors the evolution of market conditions using two costly signals that it observes in each period: a review signal, which is used to decide if the policy has become obsolete such that it is worthwhile to pay the fixed cost to redesign it, and a price signal, which is used to decide which price, from the menu of prices specified by the current policy, the firm should charge in the period. We can interpret these two signals as the information acquired by two different managers in the firm: a manager at headquarters, monitoring the overall performance of the firm's policy, and a "floor" manager, monitoring the day-to-day fluctuations that might warrant temporary price adjustments. The cost of each signal is linear in Shannon's (1948) mutual information between the signal and the state of the economy. Mutual information measures the reduction in uncertainty about the state of the economy achieved by an optimally designed signal. Uncertainty is measured by entropy, and the signal is optimal for the decision that is based on its information content. More informative signals—which reduce uncertainty about the optimal decisions more—are more costly. Hence, for each of its two decisions, the firm faces a trade-off between closely tracking the action warranted by current market conditions and economizing on information expenditure.

For tractability, there is no free memory—including regarding the passage of time—and all information, including that about past events or actions, is subject to the unit costs θ^r and θ^p for the review and pricing signals respectively. There is also no free transmission of information between the managers who make the two decisions.[26]

The Firm's Choices Given this specification, I now formalize each firm's choice of signals and define the information cost of each choice. Consider a firm undertaking a policy review in an arbitrary period t. Let $\widetilde{\omega}_t$ denote the state of the economy at the time of the review,

[25]The assumption that the review cost is fixed and yields complete information simplifies the model and may be rationalized via economies of scale in the review technology.

[26]The assumptions that information from memory is *as costly* to process as new information, and that keeping track of time is also just as costly simplify the firm's problem and the resulting optimal policy considerably. The implications of this equal-cost assumption are discussed in more detail in the appendix.

after the realization of that period's shocks. This "pre-review" state includes the current target prices as well as the history of shocks, signals and decisions through period $t-1$, for all firms in the economy. Let $\overline{V}_t(\widetilde{\omega}_t)$ be the firm's maximum attainable value, upon conducting a review, and let $V_t(\widetilde{\omega}_t)$ be the continuation value under the policy in effect at the beginning of the period. The firm's decision of whether or not to undertake a review depends on information about the difference between these two values. Extending the results of Woodford (2009), information about this difference is acquired in the form of a binary signal indicating whether or not to review the policy. Such a signal that directly indicates the action to be taken ensures that the firm does not spend resources on any extraneous information that is not directly used in its decision.

Formally, the firm's **review policy** can be recast as the choice of (i) a sequence of hazard functions $\{\Lambda_{t+\tau}(\widetilde{\omega}_{t+\tau})\}_{\tau\geq 1}$, indicating the probability of a review in each future period and state of the world, and (ii) an unconditional frequency $\overline{\Lambda}_t$ with which the firm anticipates undertaking reviews over the expected life of the policy. The cost of this review policy each period is expected to be $\theta^r I^r_{t+\tau}$. Using the definition of mutual information,[27]

$$I^r_{t+\tau} = E_t\left\{I\left(\Lambda_{t+\tau}(\widetilde{\omega}_{t+\tau}), \overline{\Lambda}_t\right)\right\}, \ \forall \tau > 0, \tag{6a}$$

$$I\left(\Lambda, \overline{\Lambda}\right) = \Lambda\left[\log\Lambda - \log\overline{\Lambda}\right] + (1-\Lambda)\left[\log(1-\Lambda) - \log(1-\overline{\Lambda})\right]. \tag{6b}$$

At the time of its review, the firm also chooses its **pricing policy**, which determines how prices are set between reviews. The firm does not have to choose a single price to charge until the next review, as in Calvo or menu cost models; nor does it have to choose a pre-determined path, as in Reis (2006) or Burstein (2006). Rather, it can choose a menu of prices and a state-dependent rule for deciding which price to charge when. Let $\omega_{t+\tau}$ indicate the state that is relevant for the firm's pricing decision in period $t+\tau$, *after* firms have made their review decisions. This "post-review" state consists of the pre-review state $\widetilde{\omega}_{t+\tau}$ and the review decisions of all firms in the economy. As in the case of the review policy, the signal structure directly indicates the action to be taken, which in this case is the price to be charged. Hence, the price setting policy consists of three objects: $\mathcal{P}_t$, $\overline{f}_t(p)$, and $\{f_{t+\tau}(p|\omega_{t+\tau})\}_{\tau\geq 0}$, namely (i) the set of log prices in the menu, (ii) the unconditional discounted frequency with which the firm expects to charge the prices in this set until the next review, and (iii) the sequence of state-dependent distributions from which a price is

[27]It is convenient to exploit the symmetry of the mutual information function to write the amount of information acquired in terms of the relative entropy between the conditional and the unconditional probabilities that characterize the review policy, rather than in terms of the relative entropy between the prior and posterior state of the world, conditional on receiving the signal.

drawn in each period, conditional on the state.[28] These conditional distributions govern how closely prices track market conditions in real time. The expected cost of the information needed to implement this pricing policy is $\theta^p I^p_{t+\tau}$ in each period, where, again using the definition of mutual information, this cost is linear in the distance between the conditional and the unconditional frequencies,

$$I^p_{t+\tau} = E_t \left\{ I \left(f_{t+\tau} \left(p | \omega_{t+\tau} \right), \overline{f}_t \left(p \right) \right) \right\}, \ \forall \tau \geq 0, \tag{7a}$$

$$I \left(f, \overline{f} \right) = \sum_{p \in \mathcal{P}} f \left(p | \omega \right) \left[\log f \left(p | \omega \right) - \log \overline{f} \left(p \right) \right]. \tag{7b}$$

The Firm's Policy The policy chosen at the time of a review in some period t attains the maximum continuation value

$$\overline{V}_t \left(\widetilde{\omega}_t \right) = E_t \left\{ \Pi_t \left(\omega_t \right) + \sum_{\tau=1}^{\infty} \beta^\tau \Gamma_{t+\tau} \left(\widetilde{\omega}_{t+\tau-1} \right) W_{t+\tau} \left(\widetilde{\omega}_{t+\tau} \right) \right\}, \tag{8}$$

$$\begin{aligned} W_{t+\tau} \left(\widetilde{\omega}_{t+\tau} \right) \equiv & \left[1 - \Lambda_{t+\tau} \left(\widetilde{\omega}_{t+\tau} \right) \right] \Pi_{t+\tau} \left(\omega_{t+\tau} \right) + \Lambda_{t+\tau} \left(\widetilde{\omega}_{t+\tau} \right) \left[\overline{V}_{t+\tau} \left(\widetilde{\omega}_{t+\tau} \right) - \kappa \right] \\ & - \theta^r I \left(\Lambda_{t+\tau} \left(\widetilde{\omega}_{t+\tau} \right), \overline{\Lambda}_t \right), \end{aligned} \tag{9}$$

where $E_t \Pi_{t+\tau}(\omega_{t+\tau})$ denotes the average profit that the firm expects in $t+\tau$ given its pricing policy, net of the cost of the price signal, and where, if the policy survives to period $t+\tau$, the firm pays for the review signal in that period, and then either keeps its policy unchanged, or pays the fixed cost to review its policy and obtain the new maximum continuation value. The average per-period profit net of the cost of the pricing signal is given by[29]

$$\Pi_{t+\tau} \left(\omega_{t+\tau} \right) \equiv \sum_{p \in \mathcal{P}_t} f_{t+\tau} \left(p | \omega_{t+\tau} \right) \left\{ \pi(p; x_{t+\tau}; Y_{t+\tau}) - \theta^p \left[\log f_{t+\tau} \left(p | \omega_{t+\tau} \right) - \log \overline{f}_t \left(p \right) \right] \right\}, \tag{10}$$

and the survival probability is given by $\Gamma_{t+1} \left(\widetilde{\omega}_t \right) \equiv 1$ and, for $\tau > 1$,

$$\Gamma_{t+\tau} \left(\widetilde{\omega}_{t+\tau-1} \right) \equiv \prod_{k=1}^{\tau-1} \left[1 - \Lambda_{t+k} \left(\widetilde{\omega}_{t+k} \right) \right]. \tag{11}$$

[28]For expositional purposes and foreshadowing later results, the set of prices is countable, although nothing in the specification rules out policies featuring continuous price distributions. Note that since knowledge regarding the passage of time is assumed to be available only through the signals themselves, both the pricing decision and the review decision are defined relative to two single discounted frequencies $\overline{f}_t(p)$ and $\overline{\Lambda}_t$, indexed by the time of the review and applicable in all periods until the next review.

[29]The flow profit π defined in (3) is now redefined in terms of the log price p and the log target price x.

3.3 The Optimal Policy

Consider a firm that reviews its policy in period t. I shall index the firm's policy objects by t to indicate dependence on the aggregate state at the time the policy was reviewed. Let $\widetilde{\Phi}$ and Φ denote the relevant parts of the aggregate state—namely the joint distributions of normalized prices and targets—at the time of some subsequent review decision and pricing decision respectively. The implementation of the firm's policy depends on both idiosyncratic conditions (summarized by the firm's normalized target price) and on these distributions.

Each time it reviews its policy, the firm learns the complete state of the economy. Therefore, its decisions can be expressed as a function of the aggregate state and of idiosyncratic variables that are normalized by the state at the time of its last review. Specifically, for a firm that last reviewed its policy in period t, I define its normalized pre-review target in period $t+\tau$ as $\widetilde{y}_{t+\tau} \equiv x_{t+\tau} - x_t$. If the firm undertakes a review in period $t+\tau$, its normalized target is reset to 0; otherwise, its normalized post-review target is $y_{t+\tau} = \widetilde{y}_{t+\tau}$. Finally, I denote by $q_{t+\tau} \equiv p_{t+\tau} - x_t$ the firm's normalized price.

The Optimal Pricing Policy. *The probability that a firm that reviewed its policy in period t will charge normalized price q in aggregate state Φ when facing a normalized target y is*

$$f_t(q|y,\Phi) = \frac{\overline{f}_t(q)\exp\left\{\frac{\pi(q,y,Y(\Phi))}{\theta^p}\right\}}{\sum_{\widehat{q}\in\mathcal{Q}_t}\overline{f}_t(\widehat{q})\exp\left\{\frac{\pi(\widehat{q},y,Y(\Phi))}{\theta^p}\right\}}. \tag{12}$$

where $\mathcal{Q}_t$ is the set of prices in the menu (possibly a singleton) and $\overline{f}_t$ is the unconditional discounted frequency with which the firm expects to charge these prices until the next review.

The Optimal Review Policy. *The probability of a policy review in aggregate state $\widetilde{\Phi}$, given a normalized pre-review target $\widetilde{y}$, satisfies*

$$\frac{\Lambda_t(\widetilde{y};\widetilde{\Phi})}{1-\Lambda_t(\widetilde{y};\widetilde{\Phi})} = \frac{\overline{\Lambda}_t}{1-\overline{\Lambda}_t}\exp\left\{\frac{1}{\theta^r}\left[\overline{V}(\widetilde{\Phi}) - \kappa - V_t\left(\widetilde{y};\widetilde{\Phi}\right)\right]\right\}, \tag{13}$$

where $\overline{\Lambda}_t$ is the unconditional discounted frequency of reviews, V_t is the continuation value under the current policy, and $\overline{V}$ is the maximum continuation value upon review.

The Frequency of Reviews. *The optimal discounted frequency of policy reviews is*

$$\overline{\Lambda}_t = \frac{E_t\left\{\sum_{\tau=1}^{\infty}\beta^\tau\Gamma_t\left(\widetilde{y}^{\tau-1};\widetilde{\Phi}_{t+\tau-1}\right)\Lambda_t(\widetilde{y}_\tau;\widetilde{\Phi}_{t+\tau})\right\}}{E_t\left\{\sum_{\tau=1}^{\infty}\beta^\tau\Gamma_t\left(\widetilde{y}^{\tau-1};\widetilde{\Phi}_{t+\tau-1}\right)\right\}}, \tag{14}$$

where $\Gamma_t(\widetilde{y}^\tau;\widetilde{\Phi}_{t+\tau})$ is the probability that the policy chosen in period t continues to apply $\tau+1$

periods later, as a function of the sequences of targets and aggregate states, with $\Gamma_t(0; \widetilde{\Phi}_0) \equiv 1$ *(the policy lasts at least one period), and, for* $\tau > 0$*,*

$$\Gamma_t(\widetilde{y}^\tau; \widetilde{\Phi}_{t+\tau}) \equiv \prod_{k=1}^{\tau-1} \left[1 - \Lambda_t(\widetilde{y}_k; \widetilde{\Phi}_{t+k})\right]. \tag{15}$$

The Frequency of Prices. *The discounted frequency with which the firm expects to set the normalized price* q *is a discounted average of the conditional probabilities of charging this price under different states, weighted by the probability of reaching these states:*

$$\overline{f}_t(q) = \frac{E_t\left\{\sum\limits_{\tau=0}^{\infty} \beta^\tau \Gamma_t(\widetilde{y}^\tau; \widetilde{\Phi}_{t+\tau}) f_t(q|y_\tau, \Phi_{t+\tau})\right\}}{E_t\left\{\sum\limits_{\tau=0}^{\infty} \beta^\tau \Gamma_t(\widetilde{y}^\tau; \widetilde{\Phi}_{t+\tau})\right\}}. \tag{16}$$

The Optimal Pricing Support. *The set* Q_t *is the optimal support of the pricing policy if and only if* $Z_t(q) \leq 1$ *for all* q *and* $Z_t(q) = 1$ *for all* q *such that* $\overline{f}_t(q) > 0$*, with*

$$Z_t(q) \equiv E_t\left\{\sum_{\tau=0}^{\infty} \beta^\tau \Gamma_t(\widetilde{y}^\tau; \widetilde{\Phi}_{t+\tau}) \frac{\exp\left\{\frac{1}{\theta^p}\pi(q; y_\tau; Y(\Phi_{t+\tau}))\right\}}{\sum\limits_{q' \in Q_t} \overline{f}_t(q') \exp\left\{\frac{1}{\theta^p}\pi(q'; y_\tau; Y(\Phi_{t+\tau}))\right\}}\right\} \tag{17}$$

where the pricing policy satisfies equations (12) and (16). The associated probability distribution satisfies the fixed point $\overline{f}_t(q) = \overline{f}_t(q)\,\overline{Z}_t(q)$, $\forall q \in Q_t$.

Discussion The firm's pricing policy is defined by equations (12), (16), and (17). Since all information about the state is equally costly, the firm chooses a signalling mechanism that directly conditions on its target and on the expected output level, which are sufficient statistics for idiosyncratic and aggregate conditions. Moreover, since the firm can revise its policy, the pricing problem becomes a *static* problem over the distribution of states that it expects to face until the next review. This means that the solution inherits the properties of solutions from the static rational inattention literature. In particular, it is worth recalling three important features of an equation of the form (12): First, it exhibits partial state-dependence in that the probability of setting a particular price in a particular state is high, relative to the average probability of charging other prices in that state, if the profit from doing so is high relative to the average profit that the firm can expect in this state across all the prices on the menu. Second, the state-dependence is stochastic. Regardless of the target price, there is positive mass on *all* prices for which $\overline{f}_t(q) > 0$. This implies not only that the firm can make considerable mistakes in pricing, but also that the price may change from one period to the next even if there is no change in the fundamentals. Third, the information cost θ^p governs the degree of noise in the solution. The higher the cost, the flatter is the conditional distribution in equation (12), reducing pricing accuracy.

Equation (16) differs from the static rational inattention solution. It represents the *discounted* frequency with which the firm anticipates that it will charge different prices from its current policy, with future states mattering less for the firm's choice of a policy today.

The condition for the optimality of the support defined in equation (17) is crucial in the context of a potentially discrete solution. The value $Z_t(q)$ represents the value of charging the price q relative to the value of charging other prices $q' \in Q_t$, on average, across all possible targets y that the firm expects to encounter until its next review. The optimal support is chosen so as to equate this value across all prices in the support. Moreover, it requires that charging any other price would yield a weakly lower average value. If one can find a set of prices that satisfy the conditions in (17), then this set characterizes the uniquely optimal solution at the information cost θ^p.

The firm's review policy is defined by equations (13) and (14). The conditional probability of a policy review has the same form as the probability of a price change derived by Woodford (2009), generalizing it to the general equilibrium model with pricing policies consisting of more than one price between reviews. The review decision depends on the firm's own pre-review normalized target, and also on expected aggregate dynamics. When deciding whether or not to review its policy, the firm considers the gain from undertaking a review relative to the cost of the review κ. The dependence of the review decision on the state is imperfect: In order to economize on information costs, the optimal review signal neither rules out a review nor indicates it with certainty. When the cost of information θ^r is low, the firm can afford to acquire more information in order to make its review decision, and hence this decision becomes increasingly precise.

Equilibrium A stationary equilibrium is a set of stochastic processes $\Lambda_t(\widetilde{y};\widetilde{\Phi})$, $\overline{\Lambda}_t$, $V_t(\widetilde{y};\widetilde{\Phi})$, $\overline{V}_t$, $\overline{f}_t(q)$, $f_t(q|y,\Phi)$, Q_t that satisfy optimal firm behavior, where the relevant aggregate states are the joint distributions of pre-review and post-review prices and targets.

The steady state with idiosyncratic shocks is characterized by a set of time-invariant objects $\Lambda(\widetilde{y})$, $\overline{\Lambda}$, $V(\widetilde{y})$, $\overline{V}$, Q, $\overline{f}(q)$, and $f(q|y)$ that satisfy the conditions above for the case of zero aggregate shocks in each period, and stationary joint distributions of normalized targets and prices, pre- and post-reviews. The steady-state algorithm solves the firm's pricing policy between reviews by incorporating algorithms based on the information theory literature, namely Arimoto (1972), Blahut (1972), Csiszár (1974), and Rose (1994). Given the steady state solution, dynamics are obtained using a linear approximation to the dynamic equations of the model around the steady state, for the case of small aggregate shocks. I use the method of Reiter (2009) with Klein (2000) numerical Jacobians code. For tractability, I restrict the degree to which firms' choices of a review policy and a pricing support depend on the aggregate state and the number of firm cohorts in the equilibrium distributions.[30]

4 Numerical Results

The model is parameterized at the weekly frequency, targeting the duration, discreteness, and volatility of pricing policies identified in micro data.

4.1 Pricing Policies in the Model

Table IV shows the parameterization of the baseline single-price and multiple-price models. Most parameters are common. The parameters that determine the preferences of the representative consumer and the properties of the production function are set to values commonly used in the literature. The elasticity of substitution is $\varepsilon = 5$. The elasticity of inter-temporal substitution is $\sigma = 2.7$. The production function features decreasing returns to scale ($\gamma = 1.5$). These parameters determine the curvature and asymmetry of the profit function, which in turn affect the losses associated with mispricing. The volatility of idiosyncratic shocks and the information costs are chosen to target the frequency of policy reviews, the median shift in prices across policies, the cardinality of pricing policies, the frequency of the modal price per policy, and the frequency and size of within-policy price changes. Although these parameters are jointly optimized to target the pricing moments, I indicate in the table the statistics that are relatively more sensitive to variations in each parameter. All parameters play a role in influencing firms' incentives to acquire information, but the volatility of the shocks plays the biggest role, with small changes in the size of shocks affecting both how much the firm spends on signals and how frequently it resets its policy.

The first key numerical result is that pricing policies feature discrete prices. The solution is discrete even though the model is infinite-horizon and with Gaussian shocks. Figure 7 shows a sample price series, along with the target price that would be charged in the full information, flexible price benchmark. The shading marks the timing of policy reviews as identified by the break test. Consistent with the data, the theory generates large, transitory volatility among a small number of infrequently updated price levels. The firm's actual price tracks the target price well, especially in the medium-run, although in the short run the firm often makes large mistakes, reflecting noise in both its reviewing and pricing decisions.

[30]Costain & Nakov (2011) also use this approach to solve a general equilibrium monetary model with heterogeneous firms and state-dependent price setting. As they note, the advantage of this method is that it allows for a solution that is non-linear in the idiosyncratic shocks, while maintaining linearity in the (small) aggregate shocks.

Table IV: Baseline Parameterization

Parameter	Symbol	Values	Explanation/Target
Discount factor	β	0.9994	Annual discount rate of 3%
Elasticity of substitution	ε	5	Full info markup of 25%
Elast. of inter-temporal subst.	σ	2.7	Strategic complementarities
Inverse Frisch elasticity	ν	0	Indivisible labor
Inverse production fn. exponent	γ	1.5	Decreasing returns to scale
Fixed cost of policy review	κ	1.65; 1.8	Frequency of policy reviews
Cost of review signal	θ^r	4	Price shift across policies
Cost of price signal	θ^p	> 0.13; 0.1	Cardinality of policy
Std. dev. of idio. quality shock	σ_ξ	0.016; 0.028	Size of price changes

Note: Where there are two values, the first indicates the SPP parameterization, and the second indicates the MRP parameterization.

The second key numerical result is that the model can generate both single-price policies and multiple-price policies, depending on parameter values. In particular, there exists a finite threshold $\bar{\theta}^p$ such that for costs of the price signal below this level, the firm always chooses to acquire the pricing signal and to implement a policy with multiple prices between reviews. The level of this threshold depends on the distribution of target prices that the firm expects will be realized between reviews. This distribution is shaped by the distribution of exogenous shocks and by how quickly the chosen review policy triggers a review when the target price deviates too much from the current policy. Larger exogenous shocks or less frequent reviews that allow shocks to accumulate both result in a higher threshold and make complex pricing policies more likely.

Table V shows the model's ability to match statistics from the micro data for both SPP and MRP series.[31] For the MRP data, large shock volatility generates policies with four distinct price levels, and large price changes both within and across policies. I target more moments than there are free parameters, so the match is imperfect. Nevertheless, the model captures very well the volatility and discreteness seen in the data. The discrete solution for the firm's pricing policy yields a moderate frequency of price changes between reviews of

[31] I target statistics for the multi-rigid series excluding the price discrimination series, since the model does not feature a price discrimination motive. In the interest of space, I omit results for OFP series, whose properties are between those of SPP and MRP series; OFP pricing patterns are generated by changing the cost θ^p so as to generate a disproportionate mass at a single price in the distribution.

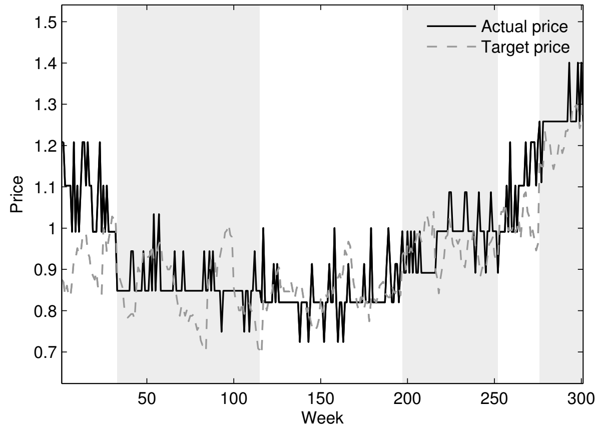


Figure 7: Simulated price series

Note: Shading marks the policy reviews identified by the break test.

39.4% versus 38.6% in the data. As in the data, policies feature one dominant rigid price, with the frequency of the modal price reaching 66% on average, versus 58% in the data.

How well do the information-constrained firms do, relative to a hypothetical firm that faces no information frictions in this economy? In the model, MRP firms achieve about 90% of the profits they would achieve if they had full information. They spend approximately 5.2% of their revenues on monitoring market conditions and updating their policies, most of which is spent on the fixed cost of policy reviews.

Since they are quite uncertain about their target price, MRP firms set prices that are 4.5 percentage points higher than the prices that would be set by fully informed firms in the same environment. Overpricing—as insurance against mistakes—reflects the fact that the firm stands ready to meet whatever demand it faces at its current price. This makes having prices that are too low much more costly than having prices that are too high, relative to the full information optimum.

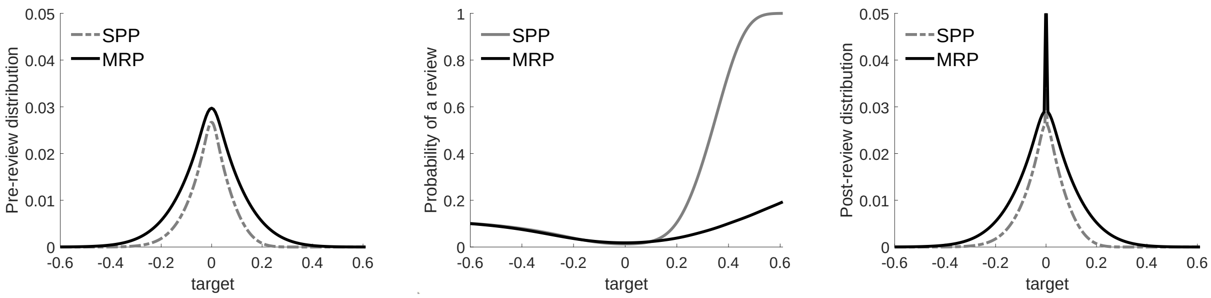
For the single-price firms, I lower the volatility of idiosyncratic shocks to match the smaller size and frequency of price changes. I also assume that redesigning single-price policies is slightly cheaper ($\kappa = 1.65$ versus 1.80 for MRP firms). Since they face less volatility in their target price, SPP firms have lower incentives to acquire information between reviews. As a result, the threshold unit cost for the price signal $\overline{\theta}^p$, which determines the desirability of having a multiple price policy, is much lower (0.13 versus 0.42 for the MRP parameterization). Overpricing is also less severe for these firms: Prices are on average 2.9 percentage points higher than the prices that would be set by fully informed firms. Lastly, profits are quite high (91% of the profits they would achieve if they had full information) even though information expenditure is less than half than that of the MRP firms.

Table V: Pricing Policies in the Model

	Single-price		Multi-rigid (non-PD)	
	Data	Model	Data	Model
Targets				
Cardinality of the pricing policy	1	1	4	4
Weekly frequency of policy reviews (%)	1.8	1.8	3.4	3.4
Shift in prices across policies (%)	8.5	8.5	10.4	10.8
Weekly freq. of modal price (%)	100	100	58.1	66.3
Weekly frequency of price changes within (%)	–	–	38.6	39.4
Size of price changes within (%)	–	–	12.2	11.1
Information expenditure				
(% of revenues)				
On reviews		1.8		3.0
On review signal		0.7		0.6
On price signal		0.0		1.6
Total info expenditure		2.5		5.2
Profits, excluding info costs (% FI)		91.2		90.0
Threshold cost $\bar{\theta}_p$ for acquiring price signal		0.13		0.42
Amount by which prices exceed FI price (%)		2.9		4.5

Note: Data versus baseline model results.

Figure 8 shows the steady state hazard functions for policy reviews for the MRP and SPP firms, and the associated steady state distributions of pre-review and post-review normalized target prices. Overall, the data favor a parameterization in which both types of firms spend relatively little on making an accurate review decision. Both hazard functions are very flat for much of the relevant state space. This implies that, all else equal, firms are slow to reset their pricing policies. Nevertheless, mispricing becomes increasingly costly when prices fall too far below the optimum. As a result, the hazard functions steepen much faster when prices fall behind, so that firms are quicker to raise prices than to cut them. The SPP hazard function displays a particularly strong steepening, since these firms cannot respond by adjusting prices between reviews.



(a) (b) (c)

Figure 8: Model outcomes

Note: Model results. (a) Distribution of pre-review targets. (b) Hazard function for policy reviews. (c) Distribution of targets post the review decision. The figures correspond to the SPP and MRP parameterizations shown in Table V.

4.2 Interdependence

To illustrate the interaction between the firm's pricing policy and its review policy, Table VI presents results for alternative parameterizations of the MRP series.[32]

Cost of Price Signal First, consider the case of a high cost of the price signal, θ^p, keeping all other parameters at the values of the MRP parameterization. The firm reduces the amount of information obtained to make its pricing decision, and instead it acquires a more precise review signal. For a high enough value of θ^p, it eliminates the price signal altogether and charges a single price between reviews. Having a more accurate timing of reviews allows the firm to undertake reviews less frequently ($\bar{\Lambda}$ declines). Hence, the firm partially makes up for its more costly price signal by spending more resources on its review policy. Nevertheless, it achieves lower profits and overprices more, since its choice of policy is more constrained.

Cost of Review Signal Next, consider an increase in θ^r, the cost of monitoring market conditions to decide whether or not to undertake a policy review. The firm now chooses a less informative review signal, which implies a flatter hazard for policy adjustment. To compensate for the increased inaccuracy in making this decision, the frequency of reviews increases, and the threshold $\overline{\theta}^p$ below which multiple-price policies are chosen also increases, making MRP policies more likely. Overall, the firm can compensate such that profits are not significantly affected.

Cost of a Review Finally, consider an increase in κ, the fixed cost of policy reviews. The firm undertakes reviews less frequently, and instead acquires more informative signals,

[32]See Alvarez & Lippi (2014) for a menu cost model discussion of how structural parameters affect pricing.

Table VI: Alternative Parameterizations for MRP Series

	Base	High θ^p	High θ^r	High κ
Policies				
Cardinality of the pricing policy	4	1	5	5
Weekly frequency of policy reviews (%)	3.4	2.8	4.4	2.8
Shift in prices across policies (%)	10.8	12.9	9.3	11.4
Information expenditure (% of revenues)				
On reviews	3.0	3.3	4.7	3.2
On review signal	0.6	1.4	1.7	0.9
On price signal	1.6	–	1.5	2.8
Total info expenditure	5.2	4.6	6.4	6.9
Profits, excluding info costs (% FI)	90.0	85.2	90.0	89.1
Amount by which prices exceed FI price (%)	4.5	4.9	4.5	5.0

Note: The first columnshows the baseline MRP parameterization. Each subsequent column considers a single parameter change: $\theta^p = 0.42$, which is the threshold information cost for multiple-price policies; $\theta^r = 20$, which generates a near-constant probability of policy reviews; and $\kappa = 3.6$, which also generates a very flat hazard function for policy reviews.

especially on pricing. It makes its review decision slightly more precise, and it designs a more complex and more accurate pricing policy. Overall, the level of spending on information increases. Profits (excluding information costs) decline, but not as much as they would if the firm had exogenously given signals. Overpricing also increases, since the firm now resets its policy less frequently, and hence there is more risk of prices becoming more stale between reviews.

Overall, the results suggest that constraints on firms' ability to design complex pricing policies may be more costly (generating lower profits and higher average prices for consumers) than having higher costs associated with the review policy, which the firm can counteract by adjusting its pricing policy between reviews. This suggests that within-policy price flexibility is a valuable way for firms to respond to shocks, a point I return to in the next section.

4.3 Discreteness

Central to obtaining a discrete solution is the shape of the distribution of target prices that the firm expects to encounter until the next review. This distribution is the key object of attention for the firm. Importantly, unlike in other rational inattention models, it is endogenous, since it is shaped by the firm's review policy which determines in which states of the world the current policy continues to apply. The review policy is more likely to trigger a review when the firm's target price has drifted far from the current menu of prices. Hence, the firm can afford to pick a small menu of prices, and then occasionally reset it. Since the profit function is asymmetric, the probability of a policy review is also asymmetric. This makes the firm more likely to reset its policy when its prices have become too low. This yields a distribution of post-review target prices whose support–while unbounded–is skewed and has negative excess kurtosis. I find numerically that these effects are strong enough to generate a discrete support for a finite cost of the price signal.

Given the optimality of a discrete support, the cost of the price signal θ^p then determines how many prices the firm chooses to charge between reviews, and how closely the probability of charging each price is tied to market conditions. Figure 9 illustrates how the pricing policy evolves in partial equilibrium, as a function of the cost of the price signal θ^p, keeping the review policy fixed. The panels plot the evolution of $Z(q)$ defined in equation (17) as a function of q, for decreasing levels of the information cost. For a high information cost, the solution yields a singleton, $\mathcal{Q} = \{\overline{q}\}$. The function Z is below 1 everywhere except at $\overline{q}$. As the information cost falls, the function Z increases for all points around $\overline{q}$. However, the growth occurs at a much faster rate in the range that will contain the new mass point. Eventually, $Z > 1$, triggering the addition of a new mass point to the optimal support. Moreover, there is no other fast-growing area over the entire range of q, such that the transition from the single-price to the multiple-price policy occurs with the growth of a *single* new mass point. This is due to the asymmetry of the problem: new mass points are added one by one to the support, spreading out over a wider and wider range of possible prices. In a setup that retains the skinny tails of the distribution of states relative to the objective function (such that discreteness remains optimal) but instead employs a symmetric objective and a symmetric distribution of states, the singleton price would "break" into two and be replaced by a price below $\overline{q}$ and a price above $\overline{q}$ simultaneously. As the cost of information is further reduced, a low price and a high price would continue to be added symmetrically. In the quadratic-normal setup, for any finite information cost, $Z(q) = 1$ for all $q \in \mathbb{R}$, as the optimal price support "breaks" to the entire real line immediately.[33]

[33]A setup in which the state is drawn from a distribution with bounded support yields a signal with a

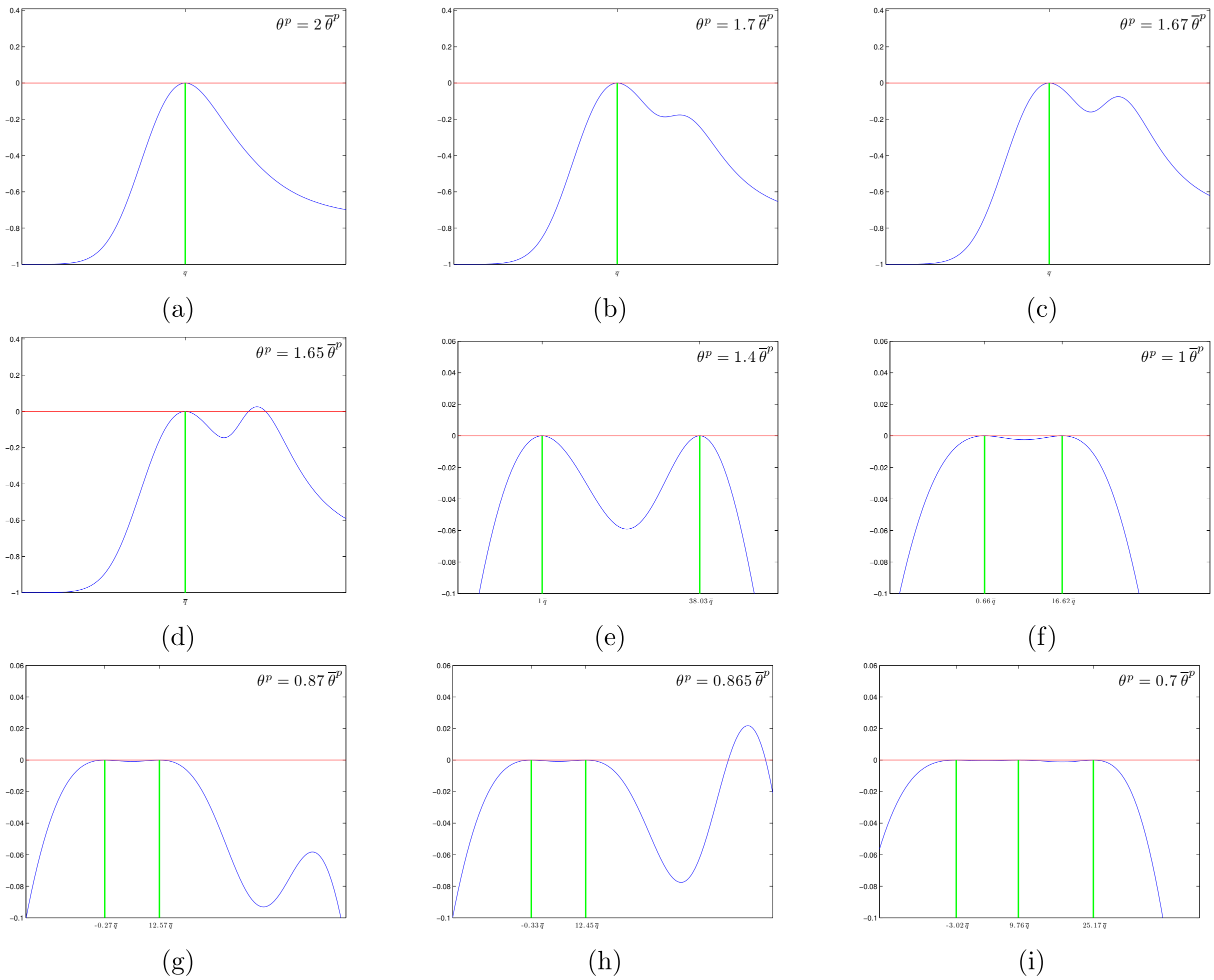


Figure 9: Growth of new mass points in the price distribution.

Note: The panels plot the function $Z(q) - 1$ as the cost of information θ^p is reduced. The points of support are shown as multiples of $\overline{q}$, the price that would be charged under the single-price policy.

Lastly, the signal endogenously allocates more attention to the regions of the state space with the potential to generate larger losses from inaccuracy. Asymmetry in the objective function implies that more attention needs to be allocated to the steeper part of the objective, since that part generates larger losses from deviating from the full-information optimum. Furthermore, depending on the distribution of shocks, attention is allocated first to the areas with more mass, and negative excess kurtosis requires less attention in the tails.

---

discrete support, regardless of the shape of the objective function, as discussed by Fix (1978), Matějka (2016) and Matějka & Sims (2010). The analysis in this paper is complementary to this work, in that I demonstrate how discreteness can arise in an infinite horizon model with Gaussian shocks.

5 Implications

What does the theory imply for the responsiveness of prices to shocks, and how does this responsiveness change if the environment becomes more volatile? In this section, I address these questions, connecting the model's micro predictions to implications for the dynamics of aggregate inflation.

5.1 Adjustment to Aggregate Shocks

Figure 10a shows how the MRP and SPP price indices respond to a contraction in aggregate nominal spending. Both series decline gradually, reflecting imprecision in pricing decisions. But MRP prices adjust faster. Since they face a higher idiosyncratic volatility, they acquire more information about market conditions and, as a result, they are also more responsive to the aggregate shock. This divergence is consistent with the patterns seen in the data during the Great Recession, when the MRP series adjusted prices more aggressively.

How much of the difference between the MRP and the SPP responses comes from the fact that MRP goods update their policies more frequently, and how much from the fact that they adjust prices between policy reviews? This split informs the question of the relevance of transitory price volatility for aggregate flexibility. The consensus that has emerged in the recent pricing literature is that such volatility does not meaningfully contribute to aggregate price flexibility. Consider filtering out the within-policy price volatility of the MRP series, and targeting only the frequency of policy reviews and the shift in prices across policies.[34] The resulting impulse response function, labeled 'Filtered' in the figure, is initially less responsive than the benchmark MRP index. But over time, it reaches and then overshoots the MRP line. The area between the two lines shows the role that within-policy price adjustment plays in responding to the aggregate shock. This dimension of adjustment is an important source of flexibility on impact and soon after the shock is realized. Most MRP firms have not yet updated their policies, but they are getting signals that they should charge the lower prices on their menus. Since these signals are partially informative, the overall MRP price index falls more than the Filtered index. But eventually, this transitory volatility actually slows down adjustment. Even after updating their policies, MRP firms continue to make mistakes in their pricing, since their price signal is imperfect. Hence, transitory price volatility has subtle effects on aggregate flexibility, flattening the IRF, and hence changing both the impact response and its subsequent persistence. I conclude that getting a truly accurate picture of how the degree of flexibility evolves over time in response

[34] This filtering out of transitory price changes requires a reparameterization of the MRP model to feature a lower volatility of idiosyncratic shocks $\sigma_\xi = 0.025$ and a lower cost of undertaking policy reviews $\kappa = 1.2$.

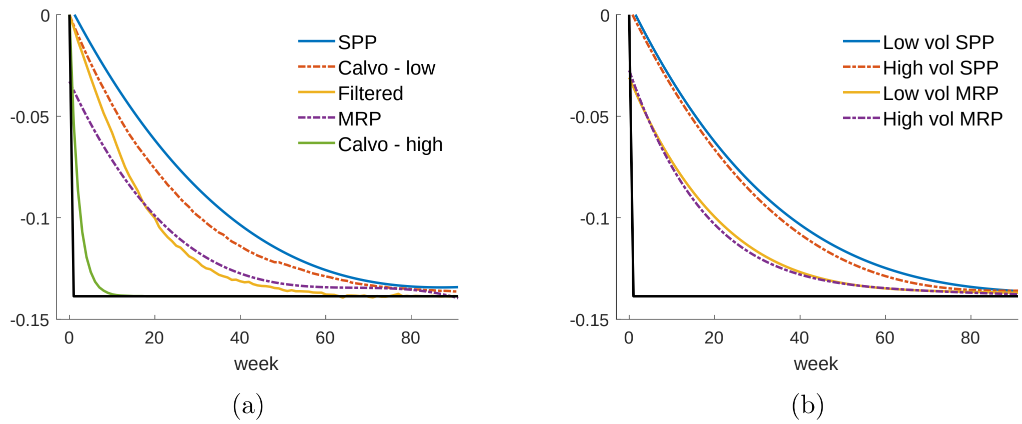


Figure 10: Impulse response functions across model specifications

Note: Model results. Panel (a) shows the impulse response functions of the price index to a negative nominal demand shock for the baseline SPP and MRP series, as well as for three alternative parameterizations: 'Filtered' shows the response of the SPP model calibrated to match the frequency of policy reviews and the shift across policies seen in the MRP data; 'Calvo-low' is the response of the standard Calvo model calibrated to match these same statistics; and 'Calvo-high' is the response of the standard Calvo model calibrated to match the frequency and size of all price changes in the MRP data. In all cases the MRP statistics are for the non-price discrimination series. Panel (b) shows the impulse response functions of the price index to the same shock in a low versus high volatility environment, for both the SPP and MRP models.

to shocks seems to require getting the dynamics of transitory price volatility right.

To put in context the responsiveness to shocks of the information-constrained firms, I consider some alternative parameterizations of the standard single-price Calvo model. First, consider a Calvo model calibrated to match the MRP frequency of policy reviews and shift in prices across policies. The resulting impulse response function is labeled 'Calvo - low' in the figure. The area between this line and the 'Filtered' line shows that the review decision of the information-constrained firm is moderately state-dependent. Alternatively, the impulse response function labeled 'Calvo - high' corresponds to a Calvo model calibrated to match the MRP frequency and size of *all* price changes. This line shows virtually no rigidity. The difference between the MRP line and this line underscores the weak relationship between the raw frequency of price changes and the degree of aggregate flexibility. This outcome reflects the noise in the firm's pricing decisions and the constraint that having a sparse menu of prices places on firms' ability to respond to shocks in real time.[35]

[35] For clarity, the figure omits the Calvo parameterization that matches the SPP frequency and size of policy adjustment. That response function is very similar to the SPP response function, highlighting the low degree of state dependence implied by the SPP hazard function. The fact that high price volatility does not necessarily imply fast adjustment to shocks has been discussed in prior work seeking to match patterns in the micro data, with prominent examples being Kehoe & Midrigan (2015) and Eichenbaum et al. (2011). However, this paper generates this result in the context of a model in which the firm chooses its policy

Table VII: The Effects of An Increase in Fundamental Volatility

	MRP series	SPP series
Change in frequency of policy reviews (%)	7.8	9.1
Change in shift across policies (%)	6.6	4.3
Change in frequency of price changes within (%)	3.0	
Change in size of price changes within (%)	1.1	
Change in total spending on information (%)	10.3	10.8
Change in profits relative to FI (ex-info) (%)	-0.7	-0.8
Change in average prices charged (%)	0.5	0.3

Note: Model results. The table shows changes in key statistics as a result of a 10% increase in volatility relative to the baseline parameterizations.

5.2 The Relationship between Volatility and Inflation

Variations in the volatility of fundamental shocks have become of increasing interest in light of the large volatility in outcomes experienced during the Great Recession. The model makes strong predictions about how volatility affects pricing policies, the aggregate price level, and its responsiveness to shocks. Table VII summarizes with a numerical illustration how the MRP and SPP policies change. Higher volatility increases the losses from having imprecise information about market conditions. As a result, it affects both the firm's review policy and its pricing policy. In a more volatile environment, spending increases on all ways of acquiring information to offset the negative effects of facing a more volatile environment. The increased uncertainty results in a large increase in the frequency of policy reviews. Conversely, the within-policy frequency and size of price changes do not change significantly. These patterns are consistent with the changes in policies that took place in the data during the Great Recession. One area where the model does not match the data concerns the shift in prices across policies. In the data, the size of the shift does not meaningfully change, whereas in the model part of the adjustment is reflected in higher shifts across policies, for both SPP and MRP series. Lastly, although the firms respond by acquiring more information, this is not enough to completely offset the negative expected effects of higher volatility, and as an additional precautionary measure, the price level also rises by half a percent.

The Great Recession was an episode market by low aggregate demand as well as heightened volatility. These forces push the firm in different directions: on the one hand, low

optimally, thereby endogenously generating the price plans postulated by Eichenbaum et al. (2011).

demand pushes the firm to reduce its prices; on the other hand, higher volatility requires setting higher prices. This tension can rationalize why inflation did not fall more during the Crisis. At the same time, it has implications for the effectiveness of monetary policy in combatting the recession. Consider the IRFs of prices to a negative demand shock when volatility is 10% higher. The model predicts that the degree of price flexibility is similar in the two economies, for both SPP and MRP series, as shown in Figure 10b. This reflects the endogenous response of information acquisition. Faced with a more uncertain environment, firms increase their information acquisition just enough to offset the higher volatility. These results contrast existing theoretical results from the menu cost model literature, where aggregate flexibility increases when volatility rises. For example, Vavra (2014) shows this result in a menu cost model with stochastic volatility.[36]

6 Conclusion

This paper argues that firms' *choice* of how much information to acquire to set prices determines aggregate price dynamics through the patterns of pricing at the micro level, and through the large heterogeneity in pricing policies across firms. Information frictions generate coarse, volatile prices that quantitatively match the patterns of price setting seen at the product level in micro data. These prices respond slowly to shocks, even though they change often. The transitory price volatility seen in the data affects the response of the price index to aggregate shocks, both in terms of the magnitude of the effect on impact and in terms of its sluggishness, though the effect is fairly modest. Finally, an increase in volatility results in a precautionary overpricing, as firms seek to protect themselves against the losses from underpricing in a more volatile environment. This rigidity in the face of a risky environment implies high monetary policy effectiveness in uncertain times. I leave for future work the question of whether cyclicality in the acquisition of information can further dampen the dynamics of inflation in response to large shocks, such as the Great Recession.

[36]The firm's ability to resolve the increased uncertainty depends on the cost function for information. In keeping with the existing rational inattention literature, I have assumed that this cost is linear in entropy reduction, but recent experimental evidence (Dean & Neligh (2017)) that the cost function for information processing might not be linear in entropy reduction. I leave this for future work.

References

Alvarez, Fernando E. & Francesco Lippi (2014), "Price Setting with Menu Cost for Multi-product Firms," *Econometrica* 82(1): 89–135.

Alvarez, Fernando E., Francesco Lippi & Luigi Paciello (2011), "Optimal Price Setting with Observation and Menu Costs," *The Quarterly Journal of Economics* 126(4): 1909–1960.

Anderson, Eric, Nir Jaimovich & Duncan Simester (2015), "Price Stickiness: Empirical Evidence of the Menu Cost Model," *Review of Economics and Statistics* 97(4): 813–826.

Anderson, Eric, Benjamin A. Malin, Emi Nakamura, Jon Steinsson & Duncan Simester (2017), "Informational Rigidities and the Stickiness of Temporary Sales," *Journal of Monetary Economics* 90: 64–83.

Arimoto, Suguru (1972), "An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channels," *IEEE Transactions on Information Theory* 18(1): 14–20.

Bai, Jushan & Pierre Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica* 66(1): 47–78.

Ball, Laurence & Sandeep Mazumder (2011), "Inflation Dynamics and the Great Recession," *Brookings Papers on Economic Activity* 2011(1): 337–381.

Beraja, Martin, Erik Hurst & Juan Ospina (2018), "The Aggregate Implications of Regional Business Cycles," Working Paper, University of Chicago.

Berger, David & Joseph Vavra (2018), "Dynamics of the US price distribution," *European Economic Review* 103: 60–82.

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

Blahut, Richard E. (1972), "Computation of Channel Capacity and Rate Distortion Functions," *IEEE Transactions on Information Theory* 18(4): 460–473.

Burstein, Ariel T. (2006), "Inflation and Output Dynamics with State-Dependent Pricing Decisions," *Journal of Monetary Economics* 53(7): 1235–1257.

Caballero, Ricardo J. & Eduardo M.R.A. Engel (1993), "Microeconomic Adjustment Hazards and Aggregate Dynamics," *The Quarterly Journal of Economics* 108(2): 359–383.

Calvo, Guillermo A. (1983), "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics* 12(3): 383–398.

Carlstein, Edward (1988), "Nonparametric Change-Point Estimation," *The Annals of Statistics* 16(1): 188–197.

Chevalier, Judith A. & Anil K. Kashyap (2019), "Best Prices: Price Discrimination and Consumer Substitution," *American Economic Journal: Economic Policy* 11(1): 126–159.

Costain, James & Anton Nakov (2011), "Distributional dynamics under smoothly state-dependent pricing," *Journal of Monetary Economics* 58(6-8): 646–665.

Cover, Thomas M. & Joy A. Thomas (2006), *Elements of Information Theory*, Wiley-Interscience, New York, 2 ed.

Csiszár, Imre (1974), "On the Computation of Rate Distortion Functions," *IEEE Transactions on Information Theory* 20(1): 122–124.

Dean, Mark & Nathaniel Neligh (2017), "Experimental Tests of Rational Inattention," Tech. rep., Working Paper, Columbia University.

Del Negro, Marco, Marc P. Giannoni & Frank Schorfheide (2015), "Inflation in the Great Recession and New Keynesian Models," *American Economic Journal: Macroeconomics* 7(1): 168–196.

DellaVigna, Stefano & Matthew Gentzkow (2017), "Uniform Pricing in U.S. Retail Chains," NBER Working Paper w23996.

Deshayes, Jean & Dominique Picard (1986), "Off-Line Statistical Analysis of Change-Point Models using Nonparametric and Likelihood Methods," in *Detection of Abrupt Changes in Signals and Dynamical Systems*, pp. 103–168, Springer Berlin Heidelberg.

Eichenbaum, Martin, Nir Jaimovich & Sergio Rebelo (2011), "Reference Prices, Costs, and Nominal Rigidities," *The American Economic Review* 101(1): 234–262.

Eichenbaum, Martin, Nir Jaimovich, Sergio Rebelo & Josephine Smith (2014), "How Frequent Are Small Price Changes?" *American Economic Journal: Macroeconomics* 6(2): 137–155.

Fix, Stephen L. (1978), "Rate Distortion Functions for Squared Error Distortion Measures," in *Proceedings of the Sixteenth Annual Allerton Conference on Communication, Control, and Computing*, pp. 704–711.

Gilchrist, Simon, Raphael Schoenle, Jae Sim & Egon Zakrajšek (2017), "Inflation dynamics during the financial crisis," *American Economic Review* 107(3): 785–823.

Golosov, Mikhail & Robert E Lucas Jr (2007), "Menu costs and Phillips curves," *Journal of Political Economy* 115(2): 171–199.

Guimaraes, Bernardo & Kevin D. Sheedy (2011), "Sales and Monetary Policy," *The American Economic Review* 101(2): 844–876.

Hall, Robert E. (2011), "The Long Slump," *The American Economic Review* 101(2): 431–69.

Ilut, Cosmin L, Rosen Valchev & Nicolas Vincent (2019), "Paralyzed by fear: Rigid and discrete pricing under demand uncertainty," Tech. rep., National Bureau of Economic Research, working Paper No. 22490.

Kehoe, Patrick J. & Virgiliu Midrigan (2015), "Prices Are Sticky After All," *Journal of Monetary Economics* 75: 35–53.

Klein, Paul (2000), "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics and Control* 24(10): 1405–1423.

Klenow, Peter J. & Oleksiy Kryvtsov (2008), "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *The Quarterly Journal of Economics* 123(3): 863–904.

Klenow, Peter J. & Benjamin A. Malin (2010), "Microeconomic Evidence on Price-Setting," in *Handbook of Monetary Economics*, B.M. Friedman & M. Woodford, eds., vol. 3A, pp. 231–284, Elsevier, Amsterdam, Holland.

Klenow, Peter J. & Jonathan L. Willis (2007), "Sticky Information and Sticky Prices," *Journal of Monetary Economics* 54: 79–99.

Maćkowiak, Bartosz & Mirko Wiederholt (2015), "Business cycle dynamics under rational inattention," *The Review of Economic Studies* 82(4): 1502–1532.

Maćkowiak, Bartosz A. & Mirko Wiederholt (2009), "Optimal Sticky Prices under Rational Inattention," *The American Economic Review* 99(3): 769–803.

Mankiw, N. Gregory & Ricardo Reis (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics* 117(4): 1295–1328.

Matějka, Filip (2016), "Rationally inattentive seller: Sales and discrete pricing," *The Review of Economic Studies* 83(3): 1125–1155.

Matějka, Filip & Alisdair McKay (2015), "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review* 105(1): 272–98.

Matějka, Filip & Christopher A. Sims (2010), "Discrete Actions in Information-Constrained Tracking Problems," Working Paper, Princeton University.

Midrigan, Virgiliu (2011), "Menu Costs, Multi-Product Firms, and Aggregate Fluctuations," *Econometrica* 79(4): 1139–1180.

Nakamura, Emi & Jón Steinsson (2008), "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics* 123(4): 1415–1464.

Nakamura, Emi & Jón Steinsson (2009), "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model," *The Quarterly Journal of Economics* 125(3): 961–1013.

Nakamura, Emi & Jón Steinsson (2013), "Price Rigidity: Microeconomic Evidence and Macroeconomic Implications," *Annual Review of Economics* 5(1): 133–163.

Nakamura, Emi, Jón Steinsson, Patrick Sun & Daniel Villar (2018), "The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation," *The Quarterly Journal of Economics* 133(4): 1933–1980.

Paciello, Luigi (2012), "Monetary Policy and Price Responsiveness to Aggregate Shocks under Rational Inattention," *Journal of Money, Credit and Banking* 44(7): 1375–1399.

Paciello, Luigi & Mirko Wiederholt (2014), "Exogenous Information, Endogenous Information and Optimal Monetary Policy," *The Review of Economic Studies* 81(1): 356–388.

Pasten, Ernesto & Raphael Schoenle (2016), "Rational Inattention, Multi-Product Firms and the Neutrality of Money," .

Reis, Ricardo (2006), "Inattentive Producers," *The Review of Economic Studies* 73(3): 793–821.

Reiter, Michael (2009), "Solving heterogeneous-agent models by projection and perturbation," *Journal of Economic Dynamics and Control* 33(3): 649–665.

Rose, Kenneth (1994), "A Mapping Approach to Rate-Distortion Computation and Analysis," *IEEE Transactions on Information Theory* 40(6): 1939–1952.

Shannon, Claude E. (1948), "A Mathematical Theory of Communication," *Bell System Technical Journal* 27: 379–423 and 623–656.

Shannon, Claude E. (1959), "Coding Theorems for a Discrete Source with a Fidelity Criterion," *Institute of Radio Engineers National Convention Record* 4: 142–163.

Sims, Christopher A. (2003), "Implications of Rational Inattention," *Journal of Monetary Economics* 50(3): 665–690.

Sims, Christopher A. (2006), "Rational Inattention: Beyond the Linear-Quadratic Case," *The American Economic Review* 96(2): 158–163.

Sims, Christopher A. (2010), "Rational Inattention and Monetary Economics," in *Handbook of Monetary Economics*, B.M. Friedman & M. Woodford, eds., vol. 3A, pp. 155–182, Elsevier, Amsterdam, Holland.

Stevens, Luminita (2011), "Pricing Regimes in Disaggregated Data," Working Paper, Columbia University.

Stevens, Luminita (2012), "Price Adjustment in a Model with Multiple Price Policies," Working Paper, University of Maryland.

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

Watson, Mark W (2014), "Inflation persistence, the NAIRU, and the great recession," *American Economic Review* 104(5): 31–36.

Woodford, Michael (2002), "Imperfect Common Knowledge and The Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, J. Stiglitz P. Aghion, R. Frydman & M. Woodford, eds., pp. 25–58, Princeton University Press, Princeton, NJ.

Woodford, Michael (2008), "Inattention as a Source of Randomized Discrete Adjustment," Working Paper, Columbia University.

Woodford, Michael (2009), "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics* 56(S): 100–124.

Online Appendix for Coarse Pricing Policies

Luminita Stevens

University of Maryland

June 2019

A Empirical Method

This section details the empirical method, its robustness across data generating processes, and the comparison with filters that seek to identify changes in regular or reference prices, rather than changes in pricing policies.

A.1 The Break Test

Test Statistic

Let $\{p_1, p_2, ..., p_n\}$ be a sequence of n price observations and define T_n as the set of all possible break points, $T_n \equiv \{t|1 \leq t < n\}$. For every hypothetical break point $t \in T_n$, the Kolmogorov-Smirnov distance between the samples $\{p_1, p_2, ..., p_t\}$ and $\{p_{t+1}, p_{t+2}, ..., p_n\}$ is

$$D_n(t) \equiv \sup_p |F_{1,t}(p) - G_{t+1,n}(p)|,$$

where $F_{1,t}$ and $G_{t+1,n}$ are the empirical cumulative distribution functions of the two subsamples, $F_{1,t}(p) \equiv \frac{1}{t}\sum_{s=1}^{t} \mathbf{1}_{\{p_s \leq p\}}$ and $G_{t+1,n}(p) \equiv \frac{1}{n-t}\sum_{s=t+1}^{n} \mathbf{1}_{\{p_s \leq p\}}$.

Following Deshayes and Picard (1986), the test statistic to test the null hypothesis of no break on a sample of size n is

$$S_n \equiv \sqrt{n} \max_{t \in T_n} \left[\frac{t}{n} \left(\frac{n-t}{n} \right) D_n(t) \right].$$

*I am grateful to the editor and three referees for their very valuable feedback, and to Mike Woodford and Ricardo Reis for their advice and support. I would also like to thank and Boragan Aruoba, Ryan Chahrour, Doireann Fitzgerald, Christian Hellwig, Pat Kehoe, Ben Malin, Filip Matějka, Virgiliu Midrigan, Emi Nakamura, Jaromir Nosal, Ernesto Pasten, Chris Sims, Jón Steinsson, and participants at numerous seminars and conferences for helpful comments. Camilo Morales-Jimenez, Dun Jia, and Donggyu Lee provided excellent research assistance. This paper supersedes the earlier working papers Stevens (2011; 2012).
Contact: stevens@econ.umd.edu.

The normalization factor depends on the relative sizes of the two sub-samples, ensuring that the test is less likely to reject the null when one of the two sub-samples is relatively short, thus providing a less precise estimate of the population CDF for that sample.

If the null is rejected ($S_n > K$, where K is the critical value determined below), the estimate of the location of the break is given by Carlstein's (1988) statistic,

$$\tau_n \equiv \arg\max_{t \in T_n} \sqrt{\frac{t\,(n-t)}{n}} D_n(t).$$

To apply this method to series that may have multiple breaks at unknown locations, I first test for the existence of one break and estimate its location. I then apply the same process to each of the two resulting sub-series, until I fail to reject the null of no break.

Critical Value

The only aspect of the algorithm that remains to be specified is the critical value used to reject the null of no break. The existing literature on estimating breaks using Kolmogorov-Smirnov focuses on the identification of a single break. For the test of a single break at an unknown location, on observations that are drawn independently from a continuous distribution, Deshayes and Picard (1986) show that under the null hypothesis of no breaks at any $t \in T_n$,

$$S_n \rightarrow \widetilde{K} \equiv \sup_{u \in [0,1]} \sup_{v \in [0,1]} |B(u,v)|\,,$$

where $B(\cdot,\cdot)$ is the two-dimensional Brownian bridge on $[0,1]$.[1] This result provides asymptotic critical values for the test of a single break on i.i.d. data from continuous distributions. However, these values are not directly applicable to my setting: I am searching for multiple potential breaks on data that is not i.i.d., and also not drawn from continuous distributions (since prices levels are quite rigid in the data). Hence, the available asymptotic critical values are too conservative and will have low power. Starting from the critical values provided by Deshayes and Picard (1986), I determine the appropriate critical value using simulations in which I compare the results of the test with the true break locations. For simplicity, I use a single critical value across all sample sizes. The critical value (and the test statistics themselves) can be tailored to individual processes. However, good-level price series are notoriously heterogeneous, hence the specification of the test should be robust *across* different types of processes. Hence, I assume that the true data generating process for product-level prices is a mixture of different processes and I use simulations to determine a single critical value to be used across all of the simulated processes.

Simulations I simulate data as a mixture of four processes that represent commonly observed price patterns: *i* sticky prices, *ii* sticky prices with temporary deviations of variable sign, size and duration, *iii* sticky prices with transitory downward sales of variable size and

[1] For the test of a single change point at a *known* location, the normalized Kolmogorov-Smirnov statistic converges to a Brownian bridge on $[0,1]$.

duration, and *iv* sticky plans, with a small number of prices repeated over time. The frequency and size of price changes target the ranges observed in the micro data. These patterns are also consistent with recent theoretical models of price setting that have sought to model micro patterns: Calvo or simple menu cost models; the dual menu cost model of Kehoe and Midrigan (2010) ; a dynamic sticky price version of the price discrimination model by Guimaraes and Sheedy (2011); and lastly, the price plans postulated by Eichenbaum et al. (2011).

For process *i*, the simulated series is given by

$$p_{t+1} = b_{t+1} \exp\{\varepsilon_{t+1}\} + (1 - b_{t+1})\, p_t,$$

where b_t is a Bernoulli trial with probability of success $\beta \in (0,1)$, marking the transition to a new price level, and $\varepsilon_t \sim \mathcal{N}(\mu, \sigma^2)$, i.i.d. This series also corresponds to the regular price series, p^R_{t+1}, for the multiple-price processes (*ii*), (*iii*) and (*iv*). In these cases, $b_t = 1$ marks the transition to a new policy.

For process *ii*, the simulated series is given by

$$p_{t+1} = b_{t+1} \exp\{\varepsilon_{t+1}\} + (1 - b_{t+1}) \left[d_{t+1} p^R_t \exp\left(\varepsilon^T_{t+1}\right) + (1 - d_{t+1})\, p^R_t \right],$$

where d_t is a Bernoulli trial with probability of success $\delta \in (0,1)$, marking the transition to a new transitory price, which is given by a mean zero i.i.d. innovation, $\varepsilon^T_t \sim \mathcal{N}(0, \sigma^2_T)$.

For process *iii*, in addition to imposing that essentially all transitory price changes are price *cuts*, by assuming that the mean of the transitory deviations is far below that of the permanent innovations, $\varepsilon^T \sim N(\mu_T, \sigma^2_T)$, with $\mu_T + 3\sigma_T < \mu - 3\sigma$, I also allow transitory prices to last up to three periods, with the maximum length of a transitory price parameterized by l_δ, with $0 \leq l_\delta \leq 3$.

Process *iv* is generated by collapsing the simulated values from process *ii* inside each policy to three bins, such that each policy consists of only three distinct prices.

These processes are parameterized to the volatility of the prices in micro data: I target a range for the mean absolute size of price changes of $10 - 15\%$, and a range for the frequency of price changes of $10 - 25\%$. Prices in the single sticky price process change with a frequency of 3%. I eliminate from simulations all policy realizations that last only one period.

Critical Values The critical value is determined using two statistics: *positive* and *negative*. The statistic *positive* reports the number of times that the test correctly rejects the null of no break on a sub-sample, as a fraction of the number of true breaks in the simulation. A low value implies that the test is not sensitive enough, such that many breaks are not identified. Correcting this requires reducing the critical value used. The statistic *negative* reports the number of times that the test incorrectly rejects the null of no break on a sub-sample that does not contain a break, as a fraction of the number of breaks estimated by the test. A high value implies that the test yields too many false positives, hence the critical value needs to be increased. Given the iterative nature of the method, the critical value determines only how soon the algorithm stops in its search for breaks: for two critical values $K_2 > K_1$, the corresponding sets of estimated break points satisfy $T_2 \subset T_1$. Hence reducing the critical value will add new breaks, without affecting the location of the existing breaks.

Table A.1 reports the performance of the break test for different critical values, starting from the asymptotic 1% and 5% significance levels provided by Deshayes and Picard (1986). The asymptotic critical values are too conservative for this setting. Using the critical value associated with the 5% significance level, the break test correctly finds only 87% of the simulated breaks on average, across all processes. The test fails to identify relatively short policy realizations, overestimating the average policy length by six periods.

Table A.1: BREAK TEST CRITICAL VALUE

Critical value, K	0.874	0.772	0.7	**0.61**	0.6	0.5	0.4
Positive (min, % true)	83.6	85.8	87.9	**90.1**	90.2	91.9	93.7
Positive (mean, % true)	83.9	86.5	88.5	**90.8**	90.9	93.2	95.0
Negative (max, % test)	0.2	0.8	1.8	**4.7**	5.1	10.2	35.2
Negative (mean, % test)	0.1	0.3	0.7	**1.3**	1.4	4.9	12.2
Exact synch (min, % true)	91.0	90.9	90.7	**90.5**	90.4	90.4	90.3
Exact synch (mean, % true)	93.4	93.4	93.3	**93.2**	93.2	93.2	93.1
Distance to truth (mean, weeks)	2	2	2	**2**	2	2	2
Length overshoot (mean, weeks)	+7	+6	+5	**+3**	+3	−0.2	−5

Break test simulation results for different critical values, across the four simulated processes. The critical values $K = 0.874$ and $K = 0.772$ are the asymptotic 1% and 5% significance levels provided by Deshayes and Picard (1986). *Positive (% true)* is the fraction of times that the test correctly rejects the null of no break, for each simulated process, reported as the minimum and the mean across all processes. *Negative (% test)* is number of times that the test incorrectly rejects the null of no break as a fraction of the total number of breaks found by the test, reported as the maximum and the mean across all simulated processes. *Exact synch (% true)* is the number of breaks found at the exact simulated location, as a fraction of the total number of breaks in the simulation, reported as both the minimum and the average across the four processes. *Distance to truth* is the average gap (number of periods) between the test estimate of the break location and the true location, excluding exact synchronizations, using a standard nearest-neighbor method. *Length overshoot* is the average number of periods by which the test overshoots the average length of policy realizations.

Reducing the critical value improves the test's performance: $K = 0.61$ is the threshold critical value for which the *positive* rate is at least 90% for all processes, while the *negative* rate is at most 5% for all processes. On average, across all processes, this critical value yields a 91% positive rate, and only a 1% negative rate. The average length of the policy realizations identified by the break test is longer than the true average length by three periods, reflecting the weak power in identifying policies that last between two and four periods. Restricting the simulations to policies lasting at least five weeks ensures the identification of virtually all breaks and eliminates the bias in the estimated average policy length..

Upon rejection of the null, I find that the change point estimate τ_k coincides exactly with the true change point 93% of the time, and is otherwise off by two periods, on average. Importantly, neither the exact synchronization nor the average distance between the estimated breaks and the true breaks, when the two are not exactly synchronized, are meaningfully affected by the choice of the critical value, since reducing the critical value does not affect the location of existing breaks, and only adds new breaks at new locations. As a result, the synchronization between the break test and the truth is consistently at 93% and the distance to the true break is consistently two periods on average.

A.2 Comparison with Filters

I compare the break test with three existing filtering methods: a v-shaped sales filter similar to those employed by Nakamura and Steinsson (2008), the reference price filter of Eichenbaum et al. (2011), and the running mode filter of Kehoe and Midrigan (2010), which is similar to that of Chahrour (2011). These filters have been proposed to uncover stickiness in product-level pricing data once one filters out transitory price changes. For these filters, a policy is identified by the regular or reference price in effect, and a break is associated with a change in the regular or reference price.

I apply each filter and the break test to micro data from Dominick's Finer Foods stores, which is a familiar and frequently used data set, for comparability with the existing literature. For each filter parameterization, I report the following statistics: *Filter duration*, which is the median policy duration implied by the filter, obtained by computing the mean frequency of breaks in each product category, taking the median across categories, and then computing the implied duration for the product with the median frequency as $d = -1/\ln(1-f)$; *Ratio of breaks*, the ratio of the number of breaks found by the filter to the number of breaks found by the break test, computed for each series and averaged across all series; *Exact synch*, the number of breaks that are synchronized between the two methods, as a fraction of the number of breaks found by the break test (also computed for each series and then averaged across all series); *Gap between methods*, the median distance between the break points estimated by the two methods, excluding exact synchronizations.

Standard statistics of interest vary significantly across the parameterizations of the different filters. Hence, although intuitive, filters present an implementation challenge in that they allow for substantial discretion in both setting up the algorithm and choosing the parameters that determine what defines a transitory price change and how it is identified.

V-shaped Sales Filter

The v-shaped sales filters eliminate price cuts that are followed, within a pre-specified window, by a price increase to the existing regular price or to a new regular price. I implement the v-shaped sales filter following Nakamura and Steinsson (2008).

The algorithm requires four parameters: J, K, L, F. The parameter J is the period of time within which a price cut must return to a regular price in order to be considered a transitory sale. When a price cut is not followed by a return to the existing regular price, several options arise regarding how to determine the new regular price. The parameters K and L capture different potential choices about when to transition to a new regular price. The parameter $F \in \{0, 1\}$ determines whether to associate the sale with the existing regular price or with the new one.

I apply the filter with different parameterizations to Dominick's data, varying the sale window $J \in \{3, ..., 12\}$, $K, L \in \{1, ..., 12\}$ and $F \in \{0, 1\}$. The parameter J is the most important determinant of the frequency of regular price changes. The parameters K, L and F do not significantly affect the median implied duration of the regular price, but they do affect the timing of breaks, thus affecting the synchronization of the filter with the break test. For example, fixing $J = 3$ while varying the remaining parameters of the v-shaped filter increases the synchronization in the timing of breaks between the v-shaped filter and the break test from 65% to 80%. Hence I report results for parameterizations of K, L, F that yield the highest degree of synchronization between the v-shaped filter and the break test, for each value of J.

Table A.2 presents the results. Statistics vary significantly with the parameterization, with the median implied duration of regular prices increasing from 12 to 29 weeks as I increase the length of the sale window, J. Increasing J beyond 12 weeks no longer significantly impacts statistics. This sensitivity to the parameterization of the filter is quite strong, but not entirely specific to Dominick's data: Nakamura and Steinsson (2008) report that for the goods underlying the US CPI, one can obtain different values for the median frequency of price changes in monthly data. For the range of parameters they test, they find median durations ranging between 6 and 8.3 months.

The filter alone cannot provide a measure of accuracy, and hence enable us to pick the best parameterization. However, the break test is expected to have at least 90% accuracy in identifying breaks in the data, if the data is a mixture of the types of processes simulated above. Hence, I compute the synchronization of the different parameterizations of the v-shaped filter with the break test.

For most parameterizations, the v-shaped method yields shorter policy realizations compared with the break test, which yields a median implied duration of 31 weeks in Dominick's data. Divergence is primarily driven by the assumption of a fixed sale window and by the fact that the filter rules out transitory price increases. Adjusting the parameters of the v-shaped filter yields a trade-off in performance: a small sales window generates many more breaks, but improves on the synchronization in the timing of the breaks found by both methods. For

Table A.2: V-SHAPED SALES FILTER PERFORMANCE

Sales window, J (weeks)	3	7	12
Filter duration (median, weeks)	12	24	29
Ratio of breaks (mean, % break test)	360	177	155
Exact synch (mean, % break test)	80	64	58
Gap between methods (median, weeks)	3	5	7

V-shaped filter results for different parameterizations on Dominick's data. *Filter duration* is the implied duration for the median frequency of breaks across product categories. *Ratio of breaks* is number of breaks found by filter divided by number of breaks found by break test, averaged across series. *Exact synch* is number of breaks that are synchronized between the two methods divided by number of breaks found by break test, averaged across series. *Gap between methods* is median distance between the break points estimated by the two methods, excluding exact synchronizations.

example, setting $J = 3$ weeks generates 360% more breaks than the break test; but 80% of the breaks found by both methods are exactly synchronized. For breaks that are not exactly synchronized, the mean distance between the break points estimated by the two methods is three weeks. Increasing the sales window still generates 55% more breaks, but substantially reduces the method's ability to estimate the timing of breaks: synchronization between the filter and the break test falls from 80% to 58%.

In summary, the v-shaped filter presents a trade-off: a short sale window captures most of the change points identified by the break test with a relatively high degree of precision, but also generates many more additional breaks, leading to an under-estimate of the rigidity of regular prices relative to the break test; a long sale window matches the median duration of regular prices, but misses the timing of breaks.

Reference Price Filter

I next implement the reference price filter proposed by Eichenbaum et al. (2011). They split the data into calendar-based quarters and define the reference price for each quarter as the most frequently quoted price in that quarter. I consider a window length in weeks $W \in \{6, 10, 13\}$.

Table A.3 presents the results. The median implied duration of reference prices increases from 24 to 51 weeks as I increase the length of the reference window, W. For reference windows above ten weeks, I find that less than 10% of the breaks are synchronized with the

Table A.3: REFERENCE PRICE FILTER PERFORMANCE

Reference window, W (weeks)	6	10	13
Filter duration (median, weeks)	24	41	51
Ratio of breaks (mean, % break test)	168	91	72
Exact synch (mean, % break test)	13	8	5
Gap between methods (median, weeks)	2	3	3

Reference price filter results for different parameterizations on Dominick's data.

break test breaks. This low ratio is entirely due to the reference price filter imposing a fixed minimum cutoff for policy lengths, which largely assumes away the question of identifying the timing of changes in the reference price series. Since I find that the length of policies is highly variable over time, the two methods are likely to overlap exactly only by chance.

In summary, the reference price filter presents a challenge in terms of identifying the timing of policy changes.

Running Mode Price Filter

I implement the running mode filter proposed by Kehoe and Midrigan (2010), which categorizes price changes as either temporary or regular, without requiring that all temporary price changes occur from a rigid *high* price, as does the v-shaped filter. For each product, they define an artificial series called the regular price series, which is a rigid running mode of the series. Every price change that is a deviation from the regular price series is defined as temporary, whereas every price change that coincides with a change in the regular price is defined as regular. In this context, I define a policy change as a change in the regular price.

The algorithm has two key parameters: A, which determines the size of the window over which to compute the modal price, and C, a cutoff used to determine if a change in the regular price has occurred. Specifically, if within a certain window, the fraction of periods in which the price is equal to the modal price is greater than C, then the regular price is updated to be equal to the current modal price; otherwise, the regular price remains unchanged.

Table A.4 presents the results. The running mode filter is much less sensitive to parameter changes compared with the reference or v-shaped filters. The median implied duration ranges from 27 to 34 weeks across parameterizations. This filter also improves on the synchronization of breaks found by the reference price filter: at the preferred parameterization, while exact synchronization with the break test is moderately low, at 48%, the median distance between the breaks found by the filter and those found by the break test is two weeks, indicating that the two methods are fairly close.

Table A.4: RUNNING MODE FILTER PERFORMANCE

Rolling window, A (weeks)	6	10	14
Filter duration (median, weeks)	27	38	34
Ratio of breaks (mean, % break test)	144	102	117
Exact synch (mean, % break test)	52	48	42
Gap between methods (median, weeks)	2	2	2

Running mode filter results for different parameterizations on Dominick's data.

In summary, when parameterized to match the duration of policies found by the break test, the running mode filter is largely in agreement with the break test, with small differences in the timing of breaks.

Performance in Simulations

To better understand the performance of the different methods, I apply all methods to simulated data, for which the true location of the breaks is known. For each filter, I use the parameterization that yields the closest match between the filter and the break test (which turns out to be the parameterization that also yields the closest match between the filter and the truth). I use the four simulated processes described above: (*i*) Single sticky price, (*ii*) One-to-flex policies, (*iii*) Downward-flex policies, and (*iv*) Coarse multiple-price policies.

I report the following statistics: *Ratio of breaks (% truth)*, the number of breaks found by the method as a fraction of the true number of breaks in the simulation; *Exact synch (% truth)*, the number of breaks found by the method that coincide with true breaks, as a fraction of the true number of breaks; *Distance to truth*, the median distance between the break points estimated by the method and the true breaks, excluding exact synchronizations, using a standard nearest-neighbor method; *Length overshoot*, the median number of periods by which the method overestimates the length of policies.
Table A.5 reports the synchronization of the methods with the true break points. The v-shaped filter over-estimates the number of breaks, and reparameterizing it to match the frequency of breaks reduces the degree of synchronization with the actual break locations. The reference price filter misses the timing of breaks, and adjusting the parameterization cannot meaningfully improve on this dimension. The running mode filter parameterized to match the frequency of breaks obtained by the break test yields results that are close to the break test, with a high degree of synchronization at 89% versus 93% for the break test.

In summary, of all the filters, the running mode filter proposed by Kehoe and Midrigan (2010) performs best in simulations, especially once it is parameterized to yield a frequency of breaks that is close to the actual frequency in the data or in the simulation.

Table A.5: FILTER PERFORMANCE IN SIMULATIONS

Method	Break test	V-shaped	Reference	Running
Ratio of breaks (% truth)	93	186	93	94
Exact synch (% of truth)	93	59	17	89
Distance to truth (median, weeks)	2	5	3	2
Length overshoot (median, weeks)	3	-9	3	2

Break test and filter results in simulated data.

A.3 Robustness of Pricing Policies Statistics

To document the robustness of the empirical properties of the identified pricing policies I report statistics for alternative classifications, identifications and definitions of the policies.

Table A.6 presents statistics using the baseline critical value, but an alternative classification of series: All product series that have at least one MRP realization are labeled as pursuing an MRP strategy. All product series that have no such realizations, but have at least one OFP policy realization are assigned to the one-to-flex category. Finally, all products that consist entirely of single-price policies are counted in the SPP category. This categorization exhausts all product series: no series are characterized by purely flexible policies.[2]

Table A.7 presents statistics based on the rolling mode filter of Kehoe and Midrigan. Under this approach, a break to a new policy is identified by a change in the modal price charged over a rolling window. Overall, the statistics are very similar when using this filter as when using the break test, though the timing of breaks does not coincide perfectly. Although the qualitative properties of pricing policies overall and by type remain unchanged, the implied policy durations are marginally longer, the size of the shift in prices across policies is one to two percentage points smaller, and the incidence of policies in which the maximum price coincides with the modal price is lower than that identified by the break test.

Table A.8 presents statistics under the baseline break test, but where the level of observation is the policy realization for each firm-product pair. The statistics are consistent with those reported for the full series in the main text, reflecting the composition of series consisting of multiple types of policy realizations.

Table A.9 presents pricing policy statistics obtained using different critical values for identifying breaks in the data.

[2] A policy is defined as flexible if no price levels are repeated over the life of a policy realization.

Table A.6: Characteristics of Price Setting Policies: Alternative Series Classification

	All	Single-price	One-to-flex	Multi-rigid
Fraction of series (%)	100	11.0	17.0	72.0
Monthly frequency of policy changes (%)	12.2	7.4	12.4	13.6
Implied policy duration (months)	7.7	13.0	7.6	6.8
Freq. of weekly price changes within (%)	24.6	0.0	8.1	33.7
Size price changes within (%)	11.9	5.8	9.3	13.0
Size of policy shift (%)	11.3	8.3	10.7	11.7
Policy cardinality	3	1	2	3

Note: AC Nielsen Retail Scanner data, 2006-2015. *Implied policy duration* is the duration implied by the median monthly frequency of policy changes. *Frequency of weekly price changes within* is the median weekly frequency with which prices change between policy breaks. *Size of price changes within* is the absolute size of price adjustment, and is non-zero for single-price policies because the category includes series with rare deviations from the modal price, as defined in the text. *Size of policy shift* is the median absolute change in the weighted average price across policy realizations. *Policy cardinality* is the median number of unique prices charged over the life of the policy.

Table A.7: Characteristics of Price Setting Policies: Kehoe-Midrigan Filter

	All	Single-price	One-to-flex	Multi-rigid
Policy duration (months)	7.8	13.3	5.7	7.9
Policy cardinality	3	1	3	4
Policy shift (%)	9.2	8.4	9.7	9.1
Freq. price changes within (%)	24.8	0.4	15.4	36.8
Size price changes within (%)	12.4	6.4	10.6	14.0
Fraction policies with max=mode (%)	57.0	79.9	57.4	54.9
Fraction of series (%)	100	11.4	30.9	57.7

Note: AC Nielsen Retail Scanner data, 2006-2015. *Policy duration* is the median duration implied by the frequency of breaks. *Policy cardinality* is the median number of unique prices charged over the life of the policy. *Policy shift* is the median absolute change in the weighted average price across policy realizations. *Freq. price changes within* is the frequency with which prices change between policy breaks. *Size of price changes within* is non-zero for single-price policies because the category includes series with policy realizations that exhibit rare deviations from the modal price, as defined in the text. *Fraction policies with max=mode* is the fraction of policy realizations in which the maximum price is the modal price, for each type of series. *Fraction of obs.* is the fraction of observations that belong to each type of series. *MR-Discrim* reports statistics for those price discrimination multi-rigid series for which a plurality of policies have the modal price equal to the high price. Statistics are computed by taking the mean across modules in each group, and then the expenditure-weighted median across groups.

Table A.8: Characteristics of Price Setting Policies: By Policy Realization

	All	Single-price	One-to-flex	Multi-rigid
Fraction of observations (%)	100	38.5	30.0	31.6
Policy cardinality	3	1	4	7
Freq. price changes within (%)	28.1	0.0	40.0	50.0
Size price changes within (%)	9.2	4.6	9.0	12.7
Fraction policies with max=mode (%)	63.5	82.6	46.5	52.0

Note: AC Nielsen Retail Scanner Data. *Fraction of observations* is the fraction of observations that belong to each type of policy realization. *Policy cardinality* is the median number of unique prices charged over the life of the policy. *Freq. price changes within* is the frequency with which prices change between policy breaks. *Size price changes within* is non-zero for single-price policies because the category includes series with policy realizations that exhibit rare deviations from the modal price, as defined in the text. *Fraction policies with max=mode* is the fraction of policy realizations in which the maximum price is the modal price, for each type of series.

Table A.9: Characteristics of Price Setting Policies for the Full Sample: Alternative Critical Values

Critical values	0.57	0.61	0.65	0.70
Monthly frequency of policy changes (%)	14.1	12.2	10.6	8.8
Implied policy duration (months)	6.6	7.7	9.0	10.9
Freq. of weekly price changes within (%)	24.1	24.6	25.0	25.3
Size price changes within (%)	11.8	11.9	12.0	12.1
Size of policy shift (%)	11.7	11.3	11.0	10.7
Policy cardinality	3	3	3	4
Fraction of series that are SPP (%)	12.6	12.0	11.5	11.0
Fraction of series that are OFP (%)	30.7	28.5	26.9	25.2
Fraction of series that are MRP (%)	56.6	59.5	61.6	63.8

Note: AC Nielsen Retail Scanner data, 2006-2015. Baseline critical value is 0.61. A higher critical value requires a larger distance between distributions and hence identifies breaks less frequently. *Implied policy duration* is the duration implied by the median monthly frequency of policy changes. *Frequency of weekly price changes within* is the median weekly frequency with which prices change between policy breaks. *Size of price changes within* is the absolute size of price adjustment, and is non-zero for single-price policies because the category includes series with rare deviations from the modal price, as defined in the text. *Size of policy shift* is the median absolute change in the weighted average price across policy realizations. *Policy cardinality* is the median number of unique prices charged over the life of the policy.

B Proofs

B.1 The Firm's Information Choices

Assumptions on the Cost of Information

Before deriving the firm's optimal policy, I discuss the key assumptions that affect the form of the solution. I assume that the quantity of information required for both the review decision and the pricing decision is small relative to the total capacity of the decision-maker, such that each unit cost may be taken as fixed. Moreover, the same unit cost applies to all types of information that may be relevant for each decision, regardless of their degrees of complexity. The types of information that are potentially relevant to each decision include information about the current conditions, the number of periods since the last review, and the history of signals and prices since the last review. Finally, I assume that there is no free transmission of information between the agents making the two decisions. The assumption that no piece of information is available for free and that the same unit cost applies to all types of information follows Woodford (2009). It is common in dynamic rational inattention papers to assume that the entire history of past signals is available to the decision-maker for free in each period, prior to acquiring the information for that period. However, the availability of that side information is not required in the current setup, given the firm's ability to occasionally review its policy.

The assumption is that all information, including knowledge of the passage of time or past events, is subject to the same unit cost of information implies that the signal structure for each decision must be defined relative to a single frequency of reviews and a single unconditional frequency of prices, both chosen at the time of the review. If instead between reviews, the decision-maker had free access to either the entire history of past signals or the number of periods that have elapsed since the last review, the optimal policy would specify a separate frequency and decision rule for each history of prior signals, or for each period between reviews, each slightly different from the previous one. Then, in each period, the agent would draw a signal from the distribution corresponding to that period. This would complicate the optimal policy tremendously, but would likely yield little gain in terms of bringing the model closer to the data. The distribution of prices would still be discrete and policies would still be updated infrequently. Each policy would still consist of a constant set of prices, but there would be a very low probability of seeing exactly the same price level be repeated over some window of time.

The Review Policy

Let $\widetilde{\omega}_t$ denote the complete state at the time of the receipt of the review signal in period t. It includes the current realization of the permanent shock and the full history of shocks, signals, and decisions through period $t-1$. Suppose that the firm decides to review its policy. The new review policy is implemented starting in period $t+1$.

Definition 1. A *review policy*, implemented following a policy review in an arbitrary state $\widetilde{\omega}_t$ in period t, is defined by

1. $\mathcal{R}_t$, the set of possible review signals;

2. $\{\rho_{t+\tau}(r|\widetilde{\omega}_{t+\tau})\}_\tau$, the sequence of conditional probabilities for all $r \in \mathcal{R}_t$, all $\widetilde{\omega}_{t+\tau}$, and all $\tau > 0$ until the next review;

3. $\overline{\rho}_t(r)$, the unconditional frequency with which the decision-maker anticipates receiving each signal r, for all $r \in \mathcal{R}_t$, until the next review;

4. $\lambda_t : \mathcal{R}_t \to [0,1]$, the decision rule for conducting reviews, with $\lambda_t(r)$ specifying the probability of conducting a review when the signal r is received, for all $r \in \mathcal{R}_t$.

The quantity of information expected, at the time of the review, to be acquired in the implementation of this review policy in each period until the next review is

$$J^r_{t+\tau} = E_t\left\{I\left(\rho_{t+\tau}(r|\widetilde{\omega}_{t+\tau}), \overline{\rho}_t(r)\right)\right\}, \tag{B.1}$$

$$I(\rho, \overline{\rho}) \equiv \sum_{r \in \mathcal{R}_t} \rho(r|\widetilde{\omega})\left[\log \rho(r|\widetilde{\omega}) - \log \overline{\rho}(r)\right], \tag{B.2}$$

where E_t denotes expectations conditional on the state $\widetilde{\omega}_t$, on a policy review having taken place in that state, and on the policy implemented at that time. This quantity is given by the average distance between the unconditional frequency of review signals over the life of the policy, $\overline{\rho}_t$, and each conditional distribution, $\rho_{t+\tau}$.

The Pricing Policy

In each period, the price signal is received after the review decision has been made, and after the realization of the transitory shock. For any $\tau \geq 0$, let $\omega_{t+\tau}$ denote the complete state at the time of the receipt of the price signal in period $t+\tau$. As above, suppose that the firm conducts a policy review in an arbitrary state $\widetilde{\omega}_t$. The new pricing policy applies starting in period t.

Definition 2. A *pricing policy*, implemented following a policy review in an arbitrary state $\widetilde{\omega}_t$ in period t, is defined by

1. $\mathcal{S}_t$, the set of possible price signals;

2. $\{\phi_{t+\tau}(s|\omega_{t+\tau})\}_\tau$, the sequence of conditional probabilities of receiving the price signal s, for all $s \in \mathcal{S}_t$, all $\tau > 0$, and all $\omega_{t+\tau}$ until the next review;

3. $\overline{\phi}_t(s)$, the unconditional frequency with which the decision-maker anticipates receiving each price signal s, for all $s \in \mathcal{S}_t$, until the next review;

4. $\alpha_t : \mathcal{S}_t \times \mathbb{R} \to [0,1]$, the decision rule for price-setting, with $\alpha_t(p|s)$ specifying the probability of charging price $p \in \mathbb{R}$ when the price signal s is received, for all $s \in \mathcal{S}_t$.

The quantity of information expected to be acquired in the implementation of this pricing policy in each period until the next review is

$$J^p_{t+\tau} = E_t\left\{I\left(\phi_{t+\tau}(s|\omega_{t+\tau}), \overline{\phi}_t(s)\right)\right\}, \tag{B.3}$$

$$I\left(\phi,\overline{\phi}\right)=\sum_{s\in\mathcal{S}_t}\phi\left(s|\omega\right)\left[\log\phi\left(s|\omega\right)-\log\overline{\phi}\left(s\right)\right], \tag{B.4}$$

where E_t denotes expectations conditional on the state $\widetilde{\omega}_t$, on a policy review having taken place in that state, and on the policy implemented at that time.

The first three elements in each of the two definitions can be thought of as the interface between the manager and her environment, while the fourth element maps the information received through this interface into the manager's actions.

These definitions are very general. The sets of possible signals $\mathcal{R}_t$ and $\mathcal{S}_t$ can include any variables that may be useful for the decisions at hand. It is important to note that nothing in the specification rules out continuous distributions. The sets $\mathcal{R}_t$ and $\mathcal{S}_t$ have been written as countable sets only for expository purposes, but it is only once we specify the objective function and the shock processes that the optimal signals will endogenously turn out to be continuous or discrete. Likewise, the sequences of conditional probabilities, $\{\rho_{t+\tau}\}_\tau$ and $\{\phi_{t+\tau}(s|\omega_{t+\tau})\}_\tau$ can be related in an arbitrary way to the state, and these relationships can vary with each future period until the next review. The only assumption is that all information, including knowledge of the passage of time or past events, is subject to the same unit cost of information. As a result, the two signal structures must each be defined relative to a single frequency ($\overline{\rho}_t$ and $\overline{\phi}_t(s)$), and each decision-maker must apply a single decision rule (λ_t and α_t), both chosen at the time of the review.[3]

The Cheapest Signal Structure

The amount of information that is *used* by the decision-maker quantifies the reduction in uncertainty that is reflected in the agent's final decision (for example, review or do not review). Let $\Lambda_{t+\tau}(\widetilde{\omega}_{t+\tau})$ denote the probability with which the decision-maker anticipates undertaking a policy review in state $\widetilde{\omega}_{t+\tau}$ in period $t+\tau$, and let $\overline{\Lambda}_t$ denote the unconditional probability of a review across all states, under the current policy,

$$\Lambda_{t+\tau}\left(\widetilde{\omega}_{t+\tau}\right)\equiv\sum_{r\in\mathcal{R}}\lambda_t\left(r\right)\rho_{t+\tau}\left(r|\widetilde{\omega}_{t+\tau}\right), \tag{B.5}$$

$$\overline{\Lambda}_t\equiv\sum_{r\in\mathcal{R}}\lambda_t\left(r\right)\overline{\rho}_t\left(r\right). \tag{B.6}$$

Similarly, let $f_{t+\tau}(p|\omega_{t+\tau})$ denote the probability that the firm charges price p in state $\omega_{t+\tau}$ in period $t+\tau$, and let $\overline{f_t}(p)$ denote the unconditional probability that price p is charged

[3]Suppose that between reviews, the decision-maker had free access to either the entire history of past signals or the number of periods that have elapsed since the last review. In that case, the firm's policy would specify separate frequencies and decision rules for each history of prior signals, or for each period between reviews. Such a specification would complicate the model but, more importantly, it would take the model farther away from the empirical evidence, which underscores simplicity in the pricing policies chosen by firms, which most often consist of no more than three or four distinct price points.

over the life of the policy,

$$f_{t+\tau}\left(p|\omega_{t+\tau}\right) \equiv \sum_{s\in\mathcal{S}} \alpha_t\left(p|s\right)\phi_{t+\tau}\left(s|\omega_{t+\tau}\right), \tag{B.7}$$

$$\overline{f_t}\left(p\right) \equiv \sum_{s\in\mathcal{S}} \alpha_t\left(p|s\right)\overline{\phi}_t\left(s\right). \tag{B.8}$$

Lemma 1. *The most efficient policy, implemented following a policy review in an arbitrary state $\widetilde{\omega}_t$ in period t, defines $\{0,1\}$ as the set of possible review signals r, and specifies*

1. $\{\Lambda_{t+\tau}\left(\widetilde{\omega}_{t+\tau}\right)\}_\tau$*, the sequence of conditional probabilities of receiving $r=1$ (conduct a review) in state $\widetilde{\omega}_{t+\tau}$, period $t+\tau$;*
2. $\overline{\Lambda}_t$*, the anticipated unconditional frequency of reviews;*
3. $\mathcal{P}_t$*, the set of prices charged until the next review;*
4. $\{f_{t+\tau}\left(p|\omega_{t+\tau}\right)\}_\tau$*, the sequence of conditional probabilities of charging price p for all $p\in\mathcal{P}_t$, all $\tau>0$ and all $\omega_{t+\tau}$ until the next review;*
5. $\overline{f}_t\left(p\right)$*, the anticipated unconditional frequency of prices, for all $p\in\mathcal{P}_t$.*

At the time of the review, the quantities of information expected to be acquired in the implementation of this policy in each period until the next review are

$$I^r_{t+\tau} = E_t\left\{I\left(\Lambda_{t+\tau}\left(\widetilde{\omega}_{t+\tau}\right), \overline{\Lambda}_t\right)\right\}, \ \forall\tau>0, \tag{B.9}$$

$$I^p_{t+\tau} = E_t\left\{I\left(f_{t+\tau}\left(p|\omega_{t+\tau}\right), \overline{f_t}\left(p\right)\right)\right\}, \ \forall\tau\geq 0. \tag{B.10}$$

Proof. Both the review decisions and prices are distributed independently of the state, conditional on the review and price signals. By the data-processing inequality (Cover & Thomas (2006)), the relative entropy between decisions and states is weakly less than the relative entropy between signals and states. If decisions are random functions of the signals, then the inequality is strict. □

This result is not only intuitive, but it also formally defines the cheapest policy that the firm can employ in order to make its review and pricing decisions. It extends the results of Woodford (2009) to the case of pricing policies that consist of more than a single price. The quantity $I^r_{t+\tau}$ defined in equation (B.9) is the smallest quantity of information that the review manager can acquire and still make exactly the same review decisions as when acquiring $J^r_{t+\tau}$, defined in equation (B.1). Likewise, the quantity $I^p_{t+\tau}$ defined in equation (B.10) is the smallest quantity of information that the pricing manager can acquire and still make exactly the same decisions as when acquiring $J^p_{t+\tau}$, defined in equation (B.3). For instance, it would not be optimal for the policy to differentiate between states in which the decision-maker takes the same action, since by merging such signals, information costs would be reduced with no loss in the accuracy of the decision. Moreover, it would also not be efficient to randomize the decision upon receipt of the signal, since it would be cheaper to reduce the mutual information between the signal and the state instead.

B.2 The Firm's Problem

Let $\overline{V}_t(\widetilde{\omega}_t)$ denote the maximum attainable value of the firm's continuation value, looking forward from the time of a policy review in an arbitrary state $\widetilde{\omega}_t$ in period t. Let

$$\Pi_{t+\tau}(\omega_{t+\tau}) \equiv \sum\nolimits_{p\in\mathcal{P}_t} f_{t+\tau}(p|\omega_{t+\tau})\,\pi(p-x_{t+\tau}) - \theta^p I\left(f_{t+\tau}(p|\omega_{t+\tau}), \overline{f}_t(p)\right)$$

denote the firm's expected per-period profit in an arbitrary state $\omega_{t+\tau}$, $\tau \geq 0$, (after that period's transitory shock, but before receipt of the price signal), under the pricing policy in effect in that state, net of the cost of the price signal, and let

$$\Gamma_{t+\tau}(\widetilde{\omega}_{t+\tau-1}) \equiv \prod\nolimits_{k=1}^{\tau-1}\left[1-\Lambda_{t+k}(\widetilde{\omega}_{t+k})\right],$$

for $\tau > 1$, with $\Gamma_{t+1}(\widetilde{\omega}_t) \equiv 1$, denote the probability, expected at the time of the review, that the review policy chosen in period t, continues to apply τ periods later, when the history of states is given by $\widetilde{\omega}_{t+\tau-1}$. The firm's maximum continuation value at the time of the policy review is

$$\overline{V}_t(\widetilde{\omega}_t) = E_t\left\{\Pi_t(\omega_t) + \sum\nolimits_{\tau=1}^{\infty}\beta^\tau\Gamma_{t+\tau}(\widetilde{\omega}_{t+\tau-1})\,W_{t+\tau}(\omega_{t+\tau})\right\},$$

$$W_\tau(\omega_\tau) \equiv \left[1-\Lambda_\tau(\widetilde{\omega}_\tau)\right]\Pi_\tau(\omega_\tau) + \Lambda_\tau(\widetilde{\omega}_\tau)\left[\overline{V}_\tau(\widetilde{\omega}_\tau)-\kappa\right] - \theta^r I\left(\Lambda_\tau(\widetilde{\omega}_\tau), \overline{\Lambda}_t\right),$$

so that conditional on the current policy surviving all the review decisions leading to a particular state $\widetilde{\omega}_{t+\tau}$, $\tau > 0$, the firm pays the cost of the review signal. It then applies the current policy with probability $1-\Lambda_{t+\tau}(\widetilde{\omega}_{t+\tau})$, in which case it attains expected profits $\Pi_{t+\tau}(\omega_{t+\tau})$, and it undertakes a policy review with probability $\Lambda_{t+\tau}(\widetilde{\omega}_{t+\tau})$, in which case it pays the review cost κ and expects the maximum attainable value from that state onward, $\overline{V}_{t+\tau}(\widetilde{\omega}_{t+\tau})$.

Since at the time of a policy review in period t, the firm learns the complete state, $\widetilde{\omega}_t$, the firm's problem can be expressed in terms of the innovations to the state since the last review. Using the normalizations defined in the main text, and given the laws of motion for the pre-review and post-review target prices, $\widetilde{x}_t$ and x_t, the normalized variables $\widetilde{y}_\tau$, y_τ, and hence $\widetilde{\varpi}_\tau$, ϖ_τ, are distributed independently of the state $\widetilde{\omega}_t$ at the time of the policy review.

The firm's problem becomes choosing $\overline{\Lambda}$, $\{\Lambda_\tau(\widetilde{\varpi}_\tau)\}_\tau$, $\mathcal{Q}$, $\overline{f}(q)$, and $\{f_\tau(q|\varpi_\tau)\}_\tau$ to solve

$$\overline{V} = \max E\left[\Pi_0(\varpi_0) + \sum\nolimits_{\tau=1}^{\infty}\beta^\tau\Gamma_\tau(\widetilde{\varpi}_{\tau-1})\,W_\tau(\varpi_\tau)\right],$$

$$W_\tau(\varpi_\tau) \equiv \left(1-\Lambda_\tau(\widetilde{\varpi}_\tau)\right)\Pi_\tau(\varpi_\tau) + \Lambda_\tau(\widetilde{\varpi}_\tau)\left(\overline{V}-\kappa\right) - \theta^r I\left(\Lambda_\tau(\widetilde{\varpi}_\tau), \overline{\Lambda}\right),$$

$$\Pi_\tau(\varpi_\tau) \equiv \sum\nolimits_{q\in\mathcal{Q}} f_\tau(q|\varpi_\tau)\,\pi(q-y_\tau) - \theta^p I\left(f_\tau(q|\varpi_\tau), \overline{f}(q)\right),$$

$$\Gamma_\tau(\widetilde{\varpi}_{\tau-1}) \equiv \prod\nolimits_{k=1}^{\tau-1}\left[1-\Lambda_k(\widetilde{\varpi}_k)\right],\ \forall\tau > 1.$$

B.3 Solution

I obtain the solution to the firm's problem in steps, deriving each element of the optimal policy taking the other elements as given.

The Conditional Distribution of Prices

The firm's choice of an optimal pricing policy for a given review policy is reduced to the maximization of the term that directly depends on the pricing policy in the firm's objective,

$E\left\{\sum_{\tau=0}^{\infty}\beta^{\tau}\Gamma_{\tau+1}\left(\widetilde{\varpi}_{\tau}\right)\Pi_{\tau}\left(\varpi_{\tau}\right)\right\}.$

Consider the subproblem of choosing the optimal sequence of conditional price distributions, $\left\{f_{\tau}\left(q|\varpi_{\tau}\right)\right\}_{\tau}$, taking as given all other elements of the firm's policy. For each τ and each possible news state ϖ_{τ} reached under the current policy, the firm chooses the conditional distribution of normalized prices $f_{\tau}\left(q|\varpi_{\tau}\right)$ that solves

$\max_{f_{\tau}(q|\varpi_{\tau})}\Pi_{\tau}\left(\varpi_{\tau}\right) \quad \text{s.t. } \sum_{q\in Q}f_{\tau}\left(q|\varpi_{\tau}\right)=1 \text{ and } f_{\tau}\left(q|\varpi_{\tau}\right)\geq 0, \forall q\in Q.$

Let the Lagrangean multipliers on the constraints be denoted by μ and $\eta\left(q\right)$. For $f_{\tau}\left(q|\varpi\right)>0$, such that $\eta\left(q\right)=0$, differentiating with respect to $f_{\tau}\left(q|\varpi\right)$, yields

$\pi(q-y_{\tau})-\theta^{p}\left[\log f_{\tau}\left(q|\varpi_{\tau}\right)-\log\overline{f}\left(q\right)\right]-\left(\theta^{p}+\mu\right)=0.$

Rearranging, and letting $\phi\equiv\exp\left\{1+\frac{\mu}{\theta^{p}}\right\}$ yields

$f_{\tau}\left(q|\varpi_{\tau}\right)=\frac{1}{\phi}\overline{f}\left(q\right)\exp\left\{\frac{1}{\theta^{p}}\pi(q-y_{\tau})\right\}.$

Summing over q and noting that the conditional distribution only depends on the normalized post-review target price y_{τ}, and on the invariant functions π and $\overline{f}$ yields as solution the invariant distribution

$f\left(q|y_{\tau}\right)=\overline{f}\left(q\right)\frac{\exp\left\{\frac{1}{\theta^{p}}\pi(q-y_{\tau})\right\}}{\sum_{\widehat{q}\in Q}\overline{f}(\widehat{q})\exp\left\{\frac{1}{\theta^{p}}\pi(\widehat{q}-y_{\tau})\right\}}.$

Note that if $\overline{f}\left(q\right)>0$, then $f\left(q|y_{\tau}\right)>0$, such that the multiplier $\eta\left(q\right)$ is indeed zero for all q, as was assumed above.

Finally, the solution implies that the expected per-period profit is also an invariant function of the normalized target price, $\Pi_{\tau}\left(\varpi_{\tau}\right)=\Pi(y_{\tau})$.

The Hazard Function for Reviews

Consider next the firm's choice of an optimal sequence of hazard functions $\left\{\Lambda_{\tau}\left(\widetilde{\varpi}_{\tau}\right)\right\}_{\tau}$ for a given pricing policy, and further taking $\overline{\Lambda}$ as given. This problem can be given a recursive formulation by noting that the choice of the sequence $\left\{\Lambda_{\tau'}\left(\widetilde{\varpi}_{\tau'}\right)\right\}_{\tau'}$ for all $\tau'>\tau$, looking forward from an arbitrary state $\widetilde{\varpi}_{\tau}$, is independent of the choices made for periods prior to τ, or for news states that are not successors of $\widetilde{\varpi}_{\tau}$. Let $V_{\tau}\left(\widetilde{\varpi}_{\tau}\right)$ be the maximum attainable value of the firm's objective, from some period τ onwards. The firm's choice of an optimal sequence of hazard functions has the recursive form

$$V_\tau\left(\widetilde{\varpi}_\tau\right)=\max\nolimits_{\Lambda_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)} E_\tau\left\{\Pi\left(y_\tau\right)+\beta\left[\begin{array}{c}\left(1-\Lambda_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)\right)V_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)\\ +\Lambda_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)\left[\overline{V}_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)-\kappa\right]\\ -\theta^r I\left(\Lambda_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right),\overline{\Lambda}\right)\end{array}\right]\right\},$$

where E_τ integrates over all possible innovations to the state, $\widetilde{\varpi}_{\tau+1}$, that follow $\widetilde{\varpi}_\tau$ under the current review policy. For each state $\widetilde{\varpi}_{\tau+1}$, the hazard function $\Lambda_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)$ is then chosen to maximize the term in square brackets.

From the solution to the firm's optimal choice for the conditional distribution of prices, $f\left(q|y\right)$, the firm's per-period profit net of the cost of the price signal is an invariant function, $\Pi\left(y\right)$, for all y. The value $V_\tau\left(\widetilde{\varpi}_\tau\right)$ depends on the state only through the dependence of the expected profit on the value of y_τ. Since $\widetilde{y}_\tau$ is a random walk and $y_\tau=\widetilde{y}_\tau+\nu_\tau$, where ν_τ is i.i.d, then for any $\tau'\geq\tau$, the probability distributions for realizations of $\widetilde{y}_{\tau'}$ and $y_{\tau'}$ conditional on $\widetilde{\varpi}_\tau$ depend only on the value of $\widetilde{y}_\tau$.

Hence, the maximum attainable value is an invariant function that only depends on the value of $\widetilde{y}_\tau$, and the solution is of the form $\Lambda_{\tau+1}\left(\widetilde{\varpi}_{\tau+1}\right)=\Lambda\left(\widetilde{y}_{\tau+1}\right)$, where $\Lambda\left(\widetilde{y}\right)$ is a time-invariant function. The value function satisfies the fixed point equation

$$V\left(\widetilde{y}\right)=E\left\{\Pi\left(y\right)+\beta\left[\left(1-\Lambda\left(\widetilde{y}'\right)\right)V\left(\widetilde{y}'\right)+\Lambda\left(\widetilde{y}'\right)\left[\overline{V}-\kappa\right]-\theta^r I\left(\Lambda\left(\widetilde{y}'\right),\overline{\Lambda}\right)\right]\right\},$$

where E denotes expectations over all possible values $\widetilde{y}'=\widetilde{y}+\widetilde{\nu}$ and $y=\widetilde{y}+\nu$, conditional on $\widetilde{y}$, the continuation value upon conducting a review is $\overline{V}=V\left(0\right)$ and

$$\overline{V}-\kappa-V\left(\widetilde{y}_{\tau+1}\right)-\theta^r\frac{\partial I\left(\Lambda\left(\widetilde{y}\right),\overline{\Lambda}\right)}{\partial\Lambda\left(\widetilde{y}\right)}=0,\text{ with}$$

$$\frac{\partial I\left(\Lambda,\overline{\Lambda}\right)}{\partial\Lambda}=\log\frac{\Lambda}{1-\Lambda}-\log\frac{\overline{\Lambda}}{1-\overline{\Lambda}}.$$

Hence

$$\frac{\Lambda\left(\widetilde{y}\right)}{1-\Lambda\left(\widetilde{y}\right)}=\frac{\overline{\Lambda}}{1-\overline{\Lambda}}\exp\left\{\frac{1}{\theta^r}\left[\overline{V}-\kappa-V\left(\widetilde{y}\right)\right]\right\}.$$

The Frequency of Reviews

For a given pricing policy, and a given hazard function for policy reviews, and using the previous two results, the optimal frequency of reviews, $\overline{\Lambda}$, is chosen to maximize

$$E\sum\nolimits_{\tau=1}^{\infty}\beta^\tau\Gamma\left(\widetilde{y}^{\tau-1}\right)\left[\left(1-\Lambda\left(\widetilde{y}_\tau\right)\right)\Pi\left(y_\tau\right)+\Lambda\left(\widetilde{y}_\tau\right)\left[\overline{V}-\kappa\right]-\theta^r I\left(\Lambda\left(\widetilde{y}_\tau\right),\overline{\Lambda}\right)\right],$$

where $\Gamma\left(\widetilde{y}^{\tau-1}\right)\equiv\prod_{k=1}^{\tau-1}\left[1-\Lambda\left(\widetilde{y}_k\right)\right]$ for $\tau>1$, with $\Gamma\left(0\right)\equiv1$, is the policy's survival probability to period τ, which depends on the history of the pre-review target prices, $\widetilde{y}^{\tau-1}$. Holding fixed the pricing policy, the value of $\overline{V}$, and the hazard function $\Lambda\left(\widetilde{y}_\tau\right)$, this problem is reduced to minimizing the cost of the review signal over the expected life of the policy. Specifically, $\overline{\Lambda}$ solves

$\min_{\overline{\Lambda}} E \sum_{\tau=1}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau-1}\right) I\left(\Lambda\left(\widetilde{y}_{\tau}\right), \overline{\Lambda}\right)$,

where the quantity of information acquired in each period for the review decision is given by

$I\left(\Lambda, \overline{\Lambda}\right) \equiv \Lambda\left[\log \Lambda - \log \overline{\Lambda}\right] + (1-\Lambda)\left[\log(1-\Lambda) - \log\left(1-\overline{\Lambda}\right)\right].$

This minimization problem is equivalent to maximizing

$E\left\{\sum_{\tau=1}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau-1}\right)\left[\Lambda\left(\widetilde{y}_{\tau}\right) \log \overline{\Lambda} + \left(1-\Lambda\left(\widetilde{y}_{\tau}\right) \log\left(1-\overline{\Lambda}\right)\right)\right]\right\}.$

The first order condition yields

$\overline{\Lambda} = \frac{E\left\{\sum_{\tau=1}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau-1}\right) \Lambda\left(\widetilde{y}_{\tau}\right)\right\}}{E\left\{\sum_{\tau=1}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau-1}\right)\right\}}.$

The Frequency of Prices

Given the results above, the firm's pricing policy maximizes $E \sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right) \Pi\left(y_{\tau}\right)$.

Holding fixed the review policy, the support of the price signal, and the conditional price distribution, the problem of choosing the optimal anticipated frequency of prices is reduced to minimizing the total cost of the price signal over the expected life of the policy.

Specifically, $\overline{f}(q) > 0$ solves

$\min_{\overline{f}(q)} E\left\{\sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right)\left[\sum_{q \in Q} f\left(q|y_{\tau}\right)\left[\log f\left(q|y_{\tau}\right) - \log \overline{f}(q)\right]\right]\right\}$

subject to $\sum_{q \in Q} \overline{f}(q) = 1$, just as the frequency of reviews, $\overline{\Lambda}$, was shown to minimize the cost of the review signal. Forming the Lagrangian with multiplier μ, the first order condition for each q charged with positive probability yields

$E\left\{\sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right) \frac{f(q|y_{\tau})}{\overline{f}(q)}\right\} = \mu$. Summing over q yields

$\mu = E\left\{\sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right)\right\}$. Hence,

$\overline{f}(q) = \frac{E\left\{\sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right) f(q|y_{\tau})\right\}}{E\left\{\sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right)\right\}}.$

Finally, the proof that $\overline{f}$ and f specify the unique optimal pricing policy among all pricing policies with support Q follows from the strict concavity of $E \sum_{\tau=0}^{\infty} \beta^{\tau} \Gamma\left(\widetilde{y}^{\tau}\right) \Pi\left(y_{\tau}\right)$ in f and $\overline{f}$. See also Csiszar (1974) in the information theory literature.

The Optimal Support

Consider the firm's pricing objective, taking as given the review policy, $E\sum_{\tau=0}^{\infty}\beta^{\tau}\Gamma(\widetilde{y}^{\tau})\Pi(y_{\tau})$. Substituting in the optimal conditional distribution $f(q|y)$ for a given marginal $\overline{f}(q)$, the objective becomes proportional to

$$E\left\{\sum_{\tau=0}^{\infty}\beta^{\tau}\Gamma(\widetilde{y}^{\tau})\log\left[\sum_{q'\in Q}\overline{f}(q')\exp\left\{\frac{1}{\theta^{p}}\pi(q'-y_{\tau})\right\}\right]\right\}$$

subject to $\sum_{q\in Q}\overline{f}(q)=1 \quad \text{and} \quad \overline{f}(q)\geq 0$ for all q.

Let μ and $\eta(q)$ be the Lagrange multipliers on the two constraints. Differentiating with respect to $\overline{f}$ yields

$Z(q;\overline{f})-\mu+\eta(q)=0$, where

$$Z(q;\overline{f})\equiv E\left\{\sum_{\tau=0}^{\infty}\frac{\beta^{\tau}\Gamma(\widetilde{y}^{\tau})\exp\left\{\frac{1}{\theta^{p}}\pi(q-y_{\tau})\right\}}{\sum_{q'\in Q}\overline{f}(q')\exp\left\{\frac{1}{\theta^{p}}\pi(q'-y_{\tau})\right\}}\right\}.$$

For $\overline{f}(q)>0$ such that $\eta(q)=0$, multiplying by $\overline{f}(q)$ yields

$Z(q;\overline{f})\overline{f}(q)=\mu\overline{f}(q)$, and summing over q yields $\mu=1$. Hence

$$Z\left(q;\overline{f}\right)\begin{cases}\leq 1 & \text{for all } q\\ =1 & \text{if } \overline{f}(q)>0\end{cases}$$

and $\overline{f}(q)$ can be found by iterating on the fixed point $Z(q;\overline{f})\overline{f}(q)=\overline{f}(q)$.

Threshold Information Cost

Following Rose (1994), the points of support must satisfy the following necessary conditions:

$$\int G\left(y|q\right)\frac{\partial\pi(q-y)}{\partial q}dy=0,$$

$$\int G\left(y|q\right)\left[\frac{\partial^{2}\pi(q-y)}{\partial q^{2}}+\frac{1}{\theta^{p}}\left(\frac{\partial\pi(q-y)}{\partial q}\right)^{2}\right]dy\leq 0,$$

These necessary conditions imply that the single-price policy, if optimal, is defined by the price

$$\overline{q}=\arg\max_{q}\int G\left(y\right)\pi\left(q-y\right)dy.$$

and the threshold cost of the price signal that is sufficiently low such that the single-price policy is not optimal is given by

$$\overline{\theta}^{p}\equiv\frac{\int G(y)\left(\frac{\partial}{\partial q}\pi(q-y)\right)^{2}dy}{\int G(y)\left(\frac{\partial^{2}}{\partial q^{2}}\pi(q-y)\right)dy},$$ where the derivatives are evaluated at $\overline{q}$.

B.4 Law of motions for the distributions

Let $f_{t-k} \equiv f^*(\omega_{t-k})$, i.e. the distribution of prices optimally chosen by the firms that reviewed their policy in period $t-k$. Let $P_{t-k}(q_t = q|y_t = \tilde{y})$ denote the probability of q_t being q given y_t being $\tilde{y}$ that corresponds to the distribution f_{t-k}. In this definition, k denotes the age of the firm's policy in period t. For numerical implementation, assume that k is bounded above by K. A firm with the policy whose age is K is assumed to review and update its policy. Let $\tilde{\Phi}_t(q_i, \tilde{y}_j, k)$ denote the mass of firms with q_i, $\tilde{y}_j$, and k-period old policy at the beginning of the period k after the realization of the aggregate shock ν_t. Similarly, let Φ_t denote the mass of firms with q_i, $\tilde{y}_j$, and k-period old policy at the end of period t after the review decision has been made. Then,

$$
\begin{aligned}
\tilde{\Phi}_{t+1}(q_i, \tilde{y}_j, 0) &= 0 \\
\tilde{\Phi}_{t+1}(q_i, \tilde{y}_j, 1) &= P_{h_a}(\xi_{t+1} = \tilde{y}_j - \eta_{t+1}) \times P_t(q_t = q_i|y_t = 0) \times \left\{ \sum_m \sum_l \sum_n \Lambda(\tilde{y}_m, \omega_t)\tilde{\Phi}_t(q_l, \tilde{y}_m, n) \right\} \\
&\quad + P_{h_a}(\xi_{t+1} = \tilde{y}_j - \eta_{t+1}) \times P_t(q_t = q_i|y_t = 0) \times \sum_m \sum_l \{1 - \Lambda(\tilde{y}_m, \omega_t)\}\tilde{\Phi}_t(q_l, \tilde{y}_m, K) \\
\tilde{\Phi}_{t+1}(q_i, \tilde{y}_j, k) &= P_{h_a}(\xi_{t+1} = \tilde{y}_j - \tilde{y}_m - \eta_{t+1}) \\
&\quad \times \sum_m \sum_l P_{t-k+1}(q_t = q_i|y_t = \tilde{y}_m) \times \{1 - \Lambda(\tilde{y}_m, \omega_t)\}\tilde{\Phi}_t(q_l, \tilde{y}_m, k-1) \quad \text{for } k = 2, 3, ..., K
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{t+1}(q_i, \tilde{y}_j, 0) &= \mathbb{1}_{\{\tilde{y}_j = 0\}} \times P_{t+1}(q_{t+1} = q_i|y_{t+1} = 0) \times \\
&\quad \sum_m \sum_l \left\{ \sum_n \Lambda(\tilde{y}_m, \omega_t)\tilde{\Phi}_{t+1}(q_l, \tilde{y}_m, n) + \{1 - \Lambda(\tilde{y}_m, \omega_t)\}\tilde{\Phi}_{t+1}(q_l, \tilde{y}_m, K) \right\} \\
\Phi_{t+1}(q_i, \tilde{y}_j, k) &= P_{t-k+1}(q_{t+1} = q_i|y_{t+1} = \tilde{y}_j) \times \sum_l \{1 - \Lambda(\tilde{y}_j, \omega_t)\}\tilde{\Phi}_{t+1}(q_l, \tilde{y}_j, k) \quad \text{for } k = 1, 2, ..., K
\end{aligned}
$$

Finally, the distributions over q and $\tilde{y}$ are given by

$$
\begin{aligned}
\tilde{\Phi}_{t+1}(q_i, \tilde{y}_j) &= \sum_{k=0}^{K} \tilde{\Phi}_{t+1}(q_i, \tilde{y}_j, k) \\
\Phi_{t+1}(q_i, \tilde{y}_j) &= \sum_{k=0}^{K} \Phi_{t+1}(q_i, \tilde{y}_j, k)
\end{aligned}
$$

B.5 Steady State Equations

The steady state is given by the following set of equations:

$$V^{ss}(\widetilde{y}_i) = \Pi^{ss}(\widetilde{y}_i) + \beta \int\int W^{ss}(\widetilde{y}_i + \xi_i) h_\xi d\xi_i \tag{B.11}$$

$$W^{ss}(\widetilde{y}_i) = \overline{V}^{ss} - \kappa + \theta^r \log\left[\overline{\Lambda}^{ss} + (1-\overline{\Lambda}^{ss}) \exp\left\{\frac{1}{\theta^r}\left[V^{ss}(\widetilde{y}_i) - \overline{V}^{ss} + \kappa\right]\right\}\right] \tag{B.12}$$

$$\Lambda^{ss}(\widetilde{y}_i) = \frac{\frac{\overline{\Lambda}^{ss}}{1-\overline{\Lambda}^{ss}} \exp\left\{\frac{1}{\theta^r}\left[\overline{V}^{ss} - \kappa - V^{ss}(\widetilde{y}_i)\right]\right\}}{1 + \frac{\overline{\Lambda}^{ss}}{1-\overline{\Lambda}^{ss}} \exp\left\{\frac{1}{\theta^r}\left[\overline{V}^{ss} - \kappa - V^{ss}(\widetilde{y}_i)\right]\right\}} \tag{B.13}$$

$$\Pi^{ss}(y_i) = \sum_{q \in Q^{ss}} f^{ss}(q|y_i)\left[\pi(q - y_i; \widetilde{Y}^{ss}) - \theta^p \log\left(\frac{f^{ss}(q|y_i)}{\overline{f}^{ss}(q)}\right)\right] \tag{B.14}$$

$$f^{ss}(q|y_i) = \frac{\overline{f}^{ss}(q) \exp\left\{\frac{1}{\theta^p}\pi(q - y_i; \widetilde{Y}^{ss})\right\}}{\sum_{q' \in Q^{ss}} \overline{f}^{ss}(q') \exp\left\{\frac{1}{\theta^p}\pi(q' - y_i; \widetilde{Y}^{ss})\right\}} \tag{B.15}$$

$$\widetilde{Y}^{ss} = \widetilde{Y}(\Omega^{ss}) = \left\{\int e^{(1-\varepsilon)(q-y)} \Phi^{ss}(dq, dy)\right\}^{-1/(1-\varepsilon)} \tag{B.16}$$

where Φ^{ss} is the invariant steady state joint distribution of post-review prices and targets implied by the joint distribution of pre-review targets and policies in the steady state Ω^{ss},

$$\overline{\Lambda}^{ss} = \frac{J^{\Lambda,ss}(0)}{J^{1,ss}(0)}, \tag{B.17}$$

$$\tag{B.18}$$

$$\overline{f}^{ss}(q) = \frac{F^{f,ss}(q;0)}{F^{1,ss}(0)}, \tag{B.19}$$

$$\tag{B.20}$$

$$\overline{Z}^{ss}(q) = Z^{ss}(q;0), \tag{B.21}$$

where

$$J^{1,ss}(\widetilde{y}_i) = \beta \int\int \left\{1 + \left[1 - \Lambda^{ss}(\widetilde{y}_i + \xi_i)\right] J^{1,ss}(\widetilde{y}_i + \xi_i)\right\} h_\xi d\xi_i \tag{B.22}$$

$$J^{\Lambda,ss}(\widetilde{y}_i) = \beta \int\int \{\Lambda^{ss}(\widetilde{y}_i + \xi_i) + \left[1 - \Lambda^{ss}(\widetilde{y}_i + \xi_i)\right] J^{\Lambda,ss}(\widetilde{y}_i + \xi_i)\} h_\xi d\xi_i \tag{B.23}$$

$$F^{1,ss}(\widetilde{y}_i) = \int\int \left\{1 + \beta\left[1 - \Lambda^{ss}(\widetilde{y}_i + \xi_i)\right] F^{1,ss}(\widetilde{y}_i + \xi_i)\right\} h_\xi d\xi_i \tag{B.24}$$

$$F^{f,ss}(q;\widetilde{y}_i) = \int\int \{f^{ss}(q|\widetilde{y}_i) + \beta\left[1 - \Lambda^{ss}(\widetilde{y}_i + \xi_i)\right] F^{f,ss}(q; \widetilde{y}_i + \xi_i)\} h_\xi d\xi_i \tag{B.25}$$

$$Z^{ss}(q;\widetilde{y}_i) = \int\int \{X^{ss}(q;\widetilde{y}_i) + \beta\left[1 - \Lambda^{ss}(\widetilde{y}_i + \xi_i)\right] X^{ss}(q;\widetilde{y}_i + \xi_i)\} h_\xi d\xi_i \tag{B.26}$$

$$X^{ss}(q;y_i) \equiv \frac{\exp\left\{\frac{1}{\theta^p}\pi(q - y_i;\widetilde{Y}^{ss}\right\}}{\sum\limits_{q'\in Q^{ss}} \overline{f}^{ss}(q')\exp\left\{\frac{1}{\theta^p}\pi(q' - y_i;\widetilde{Y}^{ss}\right\}}. \tag{B.27}$$

C Model of Price Setting

This appendix derives the equations that characterize the model of price setting presented in the paper.

Households The problem of the representative household is standard. Inter-temporal consumer optimization yields the standard first order conditions:

$$W_t(i) = H_t(i)^{\nu} C_t^{\sigma} P_t \qquad \text{and} \qquad \frac{1}{1+i_t} = \beta E_t \left[\frac{C_{t+1}^{-\sigma} P_t}{C_t^{-\sigma} P_{t+1}} \right].$$

Intra-temporal expenditure minimization yields a demand function for each variety i,

$$C_t(i) = A_t(i)^{\varepsilon-1} P_t(i)^{-\varepsilon} P_t^{\varepsilon} C_t.$$

Firms The firm's nominal profit each period is

$$\Pi_t(i) = P_t(i) Y_t(i) - W_t(i) H_t(i).$$

Substituting the household's optimality conditions and market clearing in the firm's profit function, profit in units of marginal utility becomes

$$\pi_t(i) = Y_t^{-\sigma} \left[\left(\frac{P_t(i)}{A_t(i) P_t} \right)^{1-\varepsilon} - \left(\frac{P_t(i)}{A_t(i) P_t} \right)^{-\varepsilon\eta} Y_t^{\eta+\sigma} \right],$$

where $\eta \equiv \gamma(1+\nu)$.

Full Information Solution The first order condition with respect to $P_t(i)$ yields

$$P_t(i) = \left(\frac{\varepsilon\eta}{\varepsilon-1} \right)^{\frac{1}{\varepsilon\eta-\varepsilon+1}} Y_t^{\frac{\eta+\sigma}{\varepsilon\eta-\varepsilon+1}} P_t A_t(i).$$

Plugging this solution into the aggregate price index, the equilibrium output level in the flexible price economy is

$$Y_* = \left(\frac{\varepsilon-1}{\varepsilon\eta} \right)^{\frac{1}{\eta+\sigma}}, \quad \forall t. \tag{C.1}$$

In equilibrium, $M_t = P_t Y_t$, hence the optimal price is

$$P_t(i) = \left(\frac{\varepsilon\eta}{\varepsilon-1} \right)^{\frac{1}{\eta+\sigma}} M_t A_t(i). \tag{C.2}$$

Partial Equilibrium Suppose that the economy evolves according to the flexible price, full information equilibrium. A set of firms of measure zero are information-constrained.

Using the full-information equilibrium outcomes, the profit of a constrained firm becomes

$$\pi_t(i) = \left(\frac{\varepsilon-1}{\varepsilon\eta}\right)^{\frac{-\sigma}{\eta+\sigma}} \left[\left(\frac{P_t(i)}{X_t(i)}\right)^{1-\varepsilon} - \left(\frac{\varepsilon-1}{\varepsilon\eta}\right)\left(\frac{P_t(i)}{X_t(i)}\right)^{-\varepsilon\eta}\right],$$

where $X_t(i)$ is the optimal full-information price given by equation (C.2). Note that the profit function is maximized at $P_t(i) = X_t(i)$, hence $X_t(i)$ is also the current profit-maximizing price for the information-constrained firm in the static problem, excluding information costs. Therefore, the rationally inattentive firm would like to set a price that is as close as possible to this target, subject to the costs of acquiring information about its evolution.

Using logs, the per-period real profit of the information-constrained firm is proportional to $\pi(p_t(i) - x_t(i))$, with

$$\pi(p-x) = e^{-(\varepsilon-1)(p-x)} - \frac{\varepsilon-1}{\varepsilon\eta} e^{-\varepsilon\eta(p-x)},$$

which is the objective function introduced in the body of the paper.

Let the permanent component of the log target price be defined as

$$\widetilde{x}_t(i) \equiv \frac{1}{\eta+\sigma} \ln\left(\frac{\varepsilon\eta}{\varepsilon-1}\right) + m_t + z_t(i).$$

Then, the log target price evolves according to

$$\begin{aligned} x_t(i) &= \widetilde{x}_t(i) + \zeta_t(i), \\ \widetilde{x}_t(i) &= \widetilde{x}_{t-1}(i) + \mu_t + \xi_t(i), \end{aligned}$$

where $\zeta_t(i)$ is the transitory innovation and $\mu_t + \xi_t(i)$ is the permanent innovation. The mapping into the notation used in the main body of the paper is $\widetilde{v}_t(i) \equiv \mu_t + \xi_t(i)$, and $v_t(i) \equiv \zeta_t(i)$.

In the stationary formulation, the normalized target prices τ periods after a review has occurred, are $\widetilde{y}_0(i) = 0$, $\widetilde{y}_\tau(i) = \sum_{j=1}^{\tau}(\mu_j + \xi_j(i))$, and $y_\tau(i) \equiv \widetilde{y}_\tau(i) + \zeta_\tau(i)$. Finally, conditional on a review in period t, the information-constrained price in period $t+\tau$ is $p_{t+\tau}(i) = \widetilde{x}_t(i) + q_\tau(i)$. The per-period profit function $\pi(p_t(i) - x_t(i))$ is replaced by $\pi(q_\tau(i) - y_\tau(i))$, a function of the normalized log price and the normalized log target price.

D Algorithm

This appendix describes the numerical algorithm that solves the firm's optimal policy.

Optimal Review Algorithm For a Given Pricing Algorithm

1. Given a distribution for the permanent shock $\tilde{\nu}$, discretize $\tilde{y}$ in ny points and compute the transition probability matrix $\tilde{\pi}(\tilde{y}'|\tilde{y})$ using the Tauchen method.

2. Guess a hazard function for policy reviews $\Lambda(\tilde{y})$.

3. Compute a finite approximation to the discounted distribution of pre-review target prices over the life of the policy $\tilde{G}(\tilde{y})$.

4. Find the implied $\overline{\Lambda} = \int \Lambda(\tilde{y})\, \tilde{G}(\tilde{y})d\tilde{y}$.

5. Compute a finite approximation to the discounted distribution of post-review target prices over the life of the policy $G(y)$

6. Find the optimal pricing-policy following the algorithm described in the next section. This returns a vector of prices q^* with associated marginal and conditional distributions $\bar{f}(q^*)$ and $f(q^*|y)$.

7. Compute the expected profit function $\Pi(q - y|\tilde{y})$.

8. Iterate until convergence on the value function

$$V(q, \tilde{y}) = \Pi(q - y|\tilde{y}) + \beta \sum_{\tilde{y}'} V(q, \tilde{y}')\tilde{\pi}(\tilde{y}', \tilde{y}) \forall \tilde{y}$$

9. Compute the new hazard function,

$$\Lambda(\tilde{y})^{new} = \frac{\frac{\overline{\Lambda}}{1-\overline{\Lambda}} e^{\left\{\frac{1}{\theta^r}(V(q,0)-\kappa-V(q,\tilde{y})\right\}}}{1 + \frac{\overline{\Lambda}}{1-\overline{\Lambda}} e^{\left\{\frac{1}{\theta^r}(V(q,0)-\kappa-V(q,\tilde{y})\right\}}}$$

10. If the maximum difference between $\Lambda(\tilde{y})^{new}$ and $\Lambda(\tilde{y})$ is small enough, stop. Otherwise, update $\Lambda(\tilde{y})$ as follows and go back to step 3:

$$\Lambda(\tilde{y}) = \delta\Lambda(\tilde{y}) + (1-\delta)\Lambda(\tilde{y})^{new}, 0 < \delta \leq 1$$

Optimal Pricing Algorithm For a Given Review Policy

1. Define nq as the number of prices in the pricing policy, and $q^*_{\{nq\}}$ as the optimal pricing policy with nq different prices.

2. Find the single price policy (q^{*spp}) using the algorithm described in the next section.

3. Initialize the pricing policy. $q^*_{\{1\}} = q^{*spp}$.

4. Create a dense grid of prices q^{out}, with M equally spaced prices between $\tilde{y}^{min}$ and $\tilde{y}^{max}$, which are the minimum and maximum values for $\tilde{y}$ in the grid. Define w^{out} as the space between prices in q^{out}, and add to this grid the vector of prices $q^*_{\{nq\}}$.

5. Compute the function Z^{out} for each price $\tilde{q}$ in q^{out}:

$$Z^{out}(\tilde{q}) = \int G(y) \frac{e^{\left\{\frac{1}{\theta^p}\pi(\tilde{q},y)\right\}}}{\sum_q \bar{f}(q) e^{\left\{\frac{1}{\theta^p}\pi(q,y)\right\}}} dy$$

6. Find $\tilde{q}^*$ such that:

$$\tilde{q}^* = \arg\max_{\tilde{q}} \left\{Z^{out}(\tilde{q})\right\}$$

7. Find the closest price to $\tilde{q}^*$ in the vector $q^*_{\{nq\}}$. Call that price q^{close}

8. If the distance between q^{close} and $\tilde{q}^*$ is less than w^{out}, stop and conclude that there are no more prices in the pricing policy. Otherwise, conclude that there is another price in the pricing policy q^*, and continue to the next step.

9. Increase in one unit nq, namely $nq = nq + 1$.

10. Given nq, find the optimal pricing policy $q^*_{\{nq\}}$, $\bar{f}(q^*_{\{nq\}})$ as follows:

    (a) Given a guess for $q^*_{\{nq\}} = q^{\{n\}}$, compute the optimal marginal distributions $\bar{f}(q^{\{n\}})$ using the Blahut-Arimoto algorithm described in the last section of this appendix.

    (b) Compute:

$$\begin{aligned}
W(q^{\{n\}}) &= \int G(y|q^{\{n\}})\pi(q^{\{n\}} - y)dy \\
W'(q^{\{n\}}) &= \int G(y|q^{\{n\}}) \frac{\partial \pi(q^{\{n\}} - y)}{\partial q} dy \\
W''(q^{\{n\}}) &= \int G(y|q^{\{n\}}) \left[\frac{\partial^2 \pi(q^{\{n\}} - y)}{\partial q^2} + \frac{1}{\theta^p}\left(\frac{\partial \pi(q^{\{n\}} - y)}{\partial q}\right)^2\right] dy
\end{aligned}$$

    (c) Update your guess for $q^*_{\{nq\}}$ following Newton's algorithm:

$$q^{\{n+1\}} = q^{\{n\}} - \left[W''\left(q^{\{n\}}\right)\right]^{-1} W'\left(q^{\{n\}}\right), n \geq 1$$

    (d) If the difference between $q^{\{n+1\}}$ and $q^{\{n\}}$ is small, define $q^*_{\{nq\}} = q^{\{n+1\}}$ and stop. Otherwise, go back to step (a).

11. Go back to step 5.

Single Price Algorithm For a Given Review Policy

This algorithm assumes that the distribution $G(y)$ is known and exploits the following facts that: (i) the value function $V(q,0)$ is single peaked, and (ii) the optimal price q^* is between $[\tilde{y}^{min}, \tilde{y}^{max}]$ which are the minimum an maximum values in the grid for $\tilde{y}$.

1. Given $q^{range} = [q^{min}, q^{max}]$, define $\bar{q}$ as the mid point of q^{range}.

2. Compute the function $W(\bar{q}) = \int \pi(q-y)G(y)dy$

3. Compute the derivative $W'(\bar{q}) = \frac{\partial W(\bar{q})}{\partial q} = \int \frac{\partial \pi(q-y)}{\partial q} G(y)dy$

4. If the difference between q^{max} and q^{min}, or W', is small, $q^* = \bar{q}$. Otherwise, update q^{range} as follows and go back to step 1:

$$
\begin{aligned}
q^{range} &= [q^{min}, \bar{q}] \qquad if \quad W'(\bar{q}) < 0 \\
q^{range} &= [\bar{q}, q^{max}] \qquad if \quad W'(\bar{q}) > 0
\end{aligned}
$$

The Blahut-Arimoto Algorithm

For a given support, the optimal marginal distribution is found by iterating on

$$\overline{f}(q) = \overline{f}(q) \int \frac{\exp\left\{\frac{1}{\theta^p}\pi(q-y)\right\}}{\sum_{\widehat{q}\in\mathcal{Q}} \overline{f}(\widehat{q}) \exp\left\{\frac{1}{\theta^p}\pi(\widehat{q}-y)\right\}} G(y)\,dy.$$

For a given $\overline{f}(q)$, the conditional distribution is then given by

$$f(q|y) = \overline{f}(q) \frac{\exp\left\{\frac{1}{\theta^p}\pi(q-y)\right\}}{\sum_{\widehat{q}\in\mathcal{Q}} \overline{f}(\widehat{q}) \exp\left\{\frac{1}{\theta^p}\pi(\widehat{q}-y)\right\}}.$$

For a proof of convergence, see Csiszar (1974).

For a given grid $Q = \{q_j\}$ of size n, the algorithm proceeds as follows:

1. Initialize $\overline{f}_j^{(0)} = 1/n$, $j = 1, .., n$.

2. Compute the $n_y \times n$ matrix d whose $(ij)^{th}$ entry is given by

$$d_{ij} = \exp\left\{\frac{1}{\theta^p}\pi(q_j - y_i)\right\}.$$

3. Compute

$$D_i = \sum_{j=1}^{n} \overline{f}_j^{(k)} d_{ij}, \quad i = 1, .., n_y;$$

4. Compute

$$Z_j^{(k)} = \sum_{i=1}^{n_y} G_i \frac{d_{ij}}{D_i}, \quad j = 1, .., n;$$

$$\overline{f}_j^{(k+1)} = \overline{f}_j^{(k)} Z_j^{(k)}, \quad j = 1, .., n.$$

5. Compute

$$TU = -\sum_{j=1}^{n} \overline{f}_j^{(k+1)} \ln Z_j^{(k)}; TL = -\max_j \ln Z_j^{(k)}.$$

   If $TU - TL$ exceeds a prescribed tolerance level, go back to the beginning of step 3.

6. Compute the resulting conditional and marginal, and the associated expected profit Π and information flow I

$$f_{jk} = \overline{f}_k \frac{d_{jk}}{D_j}; \overline{f}_k = \sum_{j=1}^{n_y} f_{jk} G_j;$$

$$\Pi = \sum_{j=1}^{n_y} \sum_{k=1}^{n} \pi(q_k - y_j) f_{jk} G_j;$$

$$I = \frac{1}{\theta^p} \Pi - \sum_{j=1}^{n_y} G_j \log D_j.$$

The upper and lower triggers, TU and TL, generate, via successive iterations, a decreasing and an increasing sequence respectively, which converge to the information flow I for a given expected profit, Π, and hence information cost, θ^p (see discussion in Blahut, 1972).