

Business Cycles and Endogenous Uncertainty*

Rüdiger Bachmann[†]
RWTH Aachen U. and NBER

Giuseppe Moscarini[‡]
Yale U. and NBER

December 2012

Abstract

Recessions are times of increased uncertainty and volatility at both macro and micro levels. This robust empirical pattern is typically interpreted as the result of “uncertainty shocks” which, propagated through various frictions, impact negatively on aggregate economic activity. We explore the hypothesis that the causation runs the opposite way: negative first moment shocks induce risky behavior, which in turn raises observed cross-sectional dispersion and time series volatility of individual economic outcomes. We focus on consumer price changes. We introduce imperfect information about demand in an otherwise standard monopolistically competitive model. Each firm is not sure about the elasticity of the demand it faces, but learns it gradually from its volume of sales. Information is valuable to choose the optimal mark-up and also, due to a fixed operation cost, potentially to decide whether to exit the market. Preference shocks for each product impair learning. This demand system is fully microfounded and the model aggregates to general equilibrium. When deviating from competitors’ average prices, a firm suffers a potential profit loss, but observes the response of its revenues, thus gains information about its market power. Bad economic times are the best times to price-experiment, as the opportunity cost of price mistakes is lower and exit from the market looms larger. The model is qualitatively consistent with the observed declining hazard rate of CPI price changes. Menu costs create a tension between price stickiness and increasing hazard of price adjustment. We also present new empirical evidence from CPI microdata supporting a key prediction of the model: unusually high price volatility significantly raises the chance of exit of the item from the market, even controlling for recessions.

*We thank Veronica Guerrieri and Luigi Paciello for very helpful discussions, David Berger, Jeff Campbell, Eduardo Engel, Mikhail Golosov, Robert Shimer, Joseph Vavra and audiences at the Kiel Institute-Philadelphia FED Dec 2010 conference, Yale macro lunch, and seminars at Collegio Carlo Alberto, Chicago, Atlanta FED, Columbia, 2011 SED annual meeting in Ghent, 2011 NBER Summer Institute, 2012 AEA meetings, 2012 Workshop on Macroeconomic Dynamics at the Bank of Italy for comments. We are especially grateful to David Berger and Joseph Vavra for coding and performing the empirical work on restricted access CPI microdata at the BLS. The usual disclaimer applies.

[†]Templergraben 64, 52062 Aachen, Germany. E-mail: ruediger-bachmann@gmx.de Web: <http://www.vwlmac.rwth-aachen.de/wbaker/wb/pages/team/prof.-ruediger-bachmann-phd.php>

[‡]PO Box 208268, New Haven, CT 06520-8268, USA. E-mail: giuseppe.moscarini@yale.edu Web: <http://www.econ.yale.edu/faculty1/moscarini.htm>

1 Introduction

Bad economic times are also times of rising uncertainty. This refrain finds substantial empirical support, and resonated again throughout the Great Recession. In an influential paper, Bloom (2009) offers new evidence and interprets the causality as running from uncertainty to aggregate economic activity. This view applies not only to uncertainty about the economy’s prospects or about macroeconomic policy, as reflected in the time-varying price of risk, but also to micro-level uncertainty about the fate of individual businesses. Exogenous changes in the volatility of idiosyncratic shocks, mediated by frictions, impact mean economic activity, and may cause recessions. In Bloom’s case, the frictions are irreversibility in investment and hiring/firing, so increased uncertainty raises the real option value of waiting.

In this paper, we entertain the hypothesis of reverse causality. We explore a mechanism through which a decline in aggregate economic activity, a first moment shock, induces agents to undertake riskier activities, that result in more dispersed and volatile individual outcomes. Our view is that negative first-moment shocks to a firm’s profitability, whether idiosyncratic or aggregate, lead the firm to review its *modus operandi* and to change its strategy to survive. If our view is correct, then the root cause of aggregate fluctuations are first moment shocks, and rising uncertainty is an amplification device.

We explore this hypothesis in the context of price-setting. We focus on firms, because most of the available empirical evidence on countercyclical dispersion refers to firm-level measures of performance. We focus on prices, because this is the margin that firms can activate most quickly in response to first moment shocks, more so than production inputs, projects/business ventures, relationships with suppliers and customers.¹ Berger and Vavra (2010) document that, in the micro-data underlying the Consumer Price Index, the dispersion in log price changes is strongly countercyclical.

We revisit the standard model of monopolistic competition, and further assume that the firm does not know exactly the price elasticity of the demand it faces.² Sales do not fully reveal market power, because of unobserved preference shocks that affect the distribution of demand. Due to consumers’ preference for variety, the more the price quoted by the firm deviates from those quoted by competitors, the more sales reflect the true elasticity of demand,

¹From a leading Marketing textbook (Kotler 2000, p. 456): “Price is the marketing-mix element that produces revenue; the others produce costs. Price is also one of the most flexible elements: It can be changed quickly, unlike product features and channel commitments.”

²“Many companies try to set a price that will maximize current profits. They estimate the demand and costs associated with alternative prices and choose the price that produces maximum current profit flow, cash flow, or rate of return on investment. This strategy assumes that the firm has knowledge of its demand and cost functions; in reality, these are difficult to estimate.” (Kotler (2000), p. 462).

hence the more informative they are. In good times, firms choose similar prices, as the scale of demand is high and pricing deviations, albeit informative, are costly in terms of lost flow profits. When a firm observes a string of poor sales, it becomes increasingly pessimistic about its own market power, and starts contemplating the exit decision. At that point, the returns to price experimentation, changing prices in order to learn the demand curve, increase. The firm makes a (possibly last) attempt to understand whether its product sells well or should be discontinued. When a negative aggregate profitability shock hits the economy, the cost of experimenting and learning for the good times falls, because business is slow anyway, and the returns rise, as exit looms larger and firms “gamble for resurrection”. As firms try harder to learn their demand curves, they modify their prices drastically, generate more variable and informative sales, and further adjust prices in response to what they observe. So the dispersion and volatility of price changes rise. On the other hand, when exit becomes very likely, the firm puts less weight on continuation, when information will be useful, a reason to experiment less.

To assess all these effects we explore numerical examples. First, we present results in partial stochastic equilibrium, and verify that the dispersion of price changes indeed comoves inversely with aggregate economic activity. Second, we calculate steady state general equilibrium, when aggregate productivity is constant and optimal prices quoted are consistent with market clearing. The comparative statics effects across steady states of changes in aggregate labor productivity on price dynamics are parallel those of aggregate shocks in partial equilibrium.

This exercise reveals that a burst in price volatility is more often followed by exit. We test this prediction with CPI microdata from the BLS. We find that, controlling for the average price volatility of a CPI item, for the CPI inflation rate and average frequency of price adjustment in the sector to which the item belongs, and for the state of the business cycle, high volatility of an item’s price in the previous four months has a very strong and highly significant positive effect on the chance that the same item is taken off the market by its producer/seller.

As mentioned, the real option literature suggests that increased uncertainty discourages and delays irreversible actions, such as investing in new capital, hiring new workers, or changing prices in the presence of non convex adjustment costs. This prediction is not a foregone outcome. Larger shocks increase inertia by inducing firms to move their optimal adjustment boundaries further apart, to take advantage of the larger option value of waiting, but also push firms faster against adjustment boundaries, so the net effect is ambiguous. Bloom (2009) in partial equilibrium, and Bloom, Floetotto and Jaimovich (2009) in general equilibrium, investigate this mechanism, allowing for fixed and variable adjustment costs to labor and capital. A calibrated S,s investment model, with garden variety aggregate shocks in mean productivity and less traditional innovations in the volatility of firm-specific shocks, shows that the latter

can independently cause a drop in aggregate investment and hiring, and even a recession of plausible magnitude, followed by a sharp rebound as uncertainty is resolved.

Empirical measures of micro-level volatility tend to lag the cycle, and rise in nearly all recessions for which data are available. Aggregate asset price volatility does not lag, but is not always followed by a recession, as after the stock market crash of 1987 and the Asian crisis of 1997. This is suggestive of a causal effect from recessions to micro-level uncertainty, with a relatively quick response of the latter to the former. Our goal here is not to diminish the importance of uncertainty shocks, but rather to explore the viability of alternative explanations for countercyclical uncertainty. The mechanism stressed by Bloom (2009) may still operate: faced with fixed input adjustment costs, firms that willingly undertake risky actions on other margins, such as pricing or marketing, anticipate receiving news, good or bad, so they have an incentive to wait for those news before investing or hiring. The direction of causality, however, matters for economic policy. If agents are creating idiosyncratic risk, macro stabilization policy will also reduce micro-level uncertainty, a not necessarily desirable by-product, because it impairs learning and interferes with efficient resource reallocation.

Besides business cycles, imperfect information about the price sensitivity of demand generates two new predictions for price-setting. First, price experimentation imparts an *upward bias in prices*. A firm can learn by deviating from competitors' prices both up and down, and then observing the impact on sales, but a high price has the additional benefit of reducing physical sales, thus production costs. This seems a more general principle, which goes beyond the isoelastic utility framework. Since experimentation is more valuable when the firm is at risk of bankruptcy, that is in recessions, the upward bias is countercyclical.

The second prediction pertains to the dynamics of individual prices: the volatility of price changes declines over time. A price change away from competitors' reveals information about the demand curve. Therefore, larger subsequent changes are more likely, as more information accrues. Volatility begets volatility. As the truth is revealed, the price settles. In terms of testable predictions, large price changes away from competitors' prices should be likely followed soon by more large price changes, and the probability of a non-negligible price change declines with its duration. Our model provides a natural explanation for a *hazard rate of price changes decreasing with price duration*. When a firm adjusts a price, it learns much new information, which may cause a new adjustment. These two phenomena are exactly what Campbell and Eden (2010) find in weekly scanner data for groceries from two US cities, after controlling for temporary sales. They conclude: "Taken together, these results suggest to us that sellers extensively experiment with their prices."³ In contrast, time- and state-dependent price adjustment

³Nakamura and Steinsson (2008) also find a decreasing hazard rate of price changes in CPI items, except for

models imply a constant or increasing hazard rate. Indeed, we explore a steady state equilibrium of our model where firms face also menu costs to adjust prices. The optimal inaction region is two-dimensional, in the space of price imbalance and beliefs about the demand curve. Menu costs tilt the hazard rate of price adjustment up again, in contrast with the empirical evidence.

Our modeling contribution is the introduction of demand-curve uncertainty in the standard equilibrium model of monopolistic competition. Crucially, prices impact revenues via the true price elasticity, while the confounding shocks to the level of demand are independent of prices quoted, so the firm can control the flow of information by manipulating the price. This has a cost, forgone static profits, which is larger the more the firm deviates from the average price quoted by competitors. The strategic complementarity of prices implies strategic substitutability in information acquisition. In contrast to the literature which provides informational microfoundations to price adjustment costs (e.g., Alvarez, Lippi and Paciello 2011), in our setting the tool to acquire information is the price itself, and the cost of information acquisition is endogenous. One implication is that the hazard rate of price adjustment tends to decline in duration. The multi-armed bandit literature on the “ignorant monopolist”, that learns its own demand function through pricing decisions, offers several tractable but rather ad hoc models, which posit a primitive, ad hoc noisy demand function. See Trefler (1993) for an example. To the best of our knowledge, ours is the first general *equilibrium* model of Bayesian experimentation. It is fully microfounded, preserves nice aggregation properties, and can be solved numerically. The firm’s equilibrium filtering problem can easily be extended in several directions, for example “quality obsolescence”: the introduction of new products stochastically increases the price elasticity of demand of pre-existing products, in a way that the firm cannot directly observe and can learn only from declining sales.

The paper is organized as follows. Section 2 reviews existing research on uncertainty at cyclical frequencies, to motivate our exercise and to place it in the proper context. Section 3 describes the economy, Section 4 the problem of the household, which generates the demand curve faced by firms under imperfect information, Section 5 the firm’s optimal filtering and pricing experimentation problem, Section 6 defines general equilibrium. Section 7 illustrates numerically the individual firm problem in a stochastic economy and steady state general equilibrium. Section 8 introduces menu costs. Section 9 presents the empirical test of the model’s predictions with CPI microdata. Section 10 concludes with a description of feasible extensions of the model.

annual spikes in some services that are re-priced every 12 months. Vavra (2010) controls for unobserved item heterogeneity and finds that the hazard rate strongly declines in the first two months since last adjustment. See Alvarez, Lippi and Paciello (2011) for an extensive discussion of hazard rates.

2 Uncertainty and the business cycle

It has long been known that financial markets are more volatile when asset prices suffer, typically in bad economic times. Campbell et al. (2001) document that the volatility of individual stock prices negatively comoves with detrended GDP. Bloom (2009) shows that the VXO index of stock market volatility covaries negatively with industrial production, and positively with cross-sectional dispersion in profits and stock returns. As mentioned, the scholarly narrative interprets changes in uncertainty as random changes in the standard deviation of idiosyncratic shocks, which impact aggregate economic activity. Indeed, Bloom (2009) finds that positive innovations to the VXO index in a VAR identified by Choleski-ordering cause a sharp drop and rebound in industrial production.

More recently, the attention has moved to uncertainty in macroeconomic policies, and uncertainty thereof. Baker et. al (2012) measure uncertainty in macroeconomic policy from news media. Born and Pfeifer (2012), and Fernandez-Villaverde et al. (2011, 2012) estimate uncertainty in fundamentals, monetary and fiscal policy, and show that shocks to uncertainty about future policy have a discernible but small impact on real activity and inflation in an otherwise standard New Keynesian model. General equilibrium effects, notably the drop in the interest rate, erase most of the effects identified by Bloom (2009) in partial equilibrium. Basu and Bundick (2012) and Johanssen (2012) show that shocks to the volatility of (resp.) fundamentals, such as discount factors and technology, and of fiscal policy have a sizable impact on real activity if and only if the economy is close to the Zero Lower Bound, because the general equilibrium effect is muted. The last findings require, of course, strong nominal rigidity, and appear fragile in light of Vavra (2012)’s result that nominal rigidity and money non-neutrality are severely reduced in recessions, because the distribution of price changes also experiences a large increase in spread (Berger and Vavra 2010), so firms change prices more independently of monetary shocks.

We are interested, instead, in cyclical movements in *cross-sectional* measures of dispersion in outcomes, primarily across firms. Aggregate volatility comoves with micro-level “uncertainty”. Bloom et al. (2012) document that the dispersion of shipments across manufacturing establishments, of growth rates of sales across listed companies and across all industries are strongly countercyclical. The precise mechanism through which uncertainty shocks cause recessions relies on the type of friction that impedes continuous microeconomic adjustment.⁴

⁴An early version of this idea is Lilien (1982)’s sectorial reallocation hypothesis, based on frictions in the reallocation of employment across sectors. An increased dispersion in the size of sector-specific shocks causes temporary unemployment, as workers slowly retrain and reallocate to expanding industries. Brainard and Cutler (1993) measure sectorial shocks by means of excess returns on industry stock indices, after controlling for industry-specific betas, and indeed find evidence that the dispersion of excess returns moves over time and

Bachmann and Bayer (2011a, 2011b) use a long panel of German firms and show that the dispersion in firm-level innovations in TFP, sales and real value added is countercyclical, although the dispersion in investment rates is procyclical. Unlike much of the previous literature, they rely on direct measures of performance for a large sample of companies of all types, not just publicly traded, but also private, and not only in manufacturing, but spanning all sectors. This direct evidence indeed points to more news in firms' fundamentals during recessions. Using this evidence to calibrate a neoclassical economy with non convex adjustment costs, Bachmann and Bayer conclude that shocks to the variance of firm-level TFP innovations, if any, mildly amplify first-moment aggregate shocks, but are not an independent source of aggregate fluctuations. Bachmann, Elstner and Sims (2010) compare measures of individual business expectations, from surveys of decision makers inside the firms, with ex post business performance, to extract a measure of genuine uncertainty. They replace stock market volatility with mean square error in business forecasts and repeat Bloom (2009)'s VAR exercise. They find a negative, but very persistent, effect of uncertainty shocks on industrial output. The real option approach predicts a sharp but short-lived effect. Next, as an alternative (to ordering) identification strategy, they require that structural uncertainty shocks have no long-run effects on aggregate economic activity. They find that these shocks have no discernible impact even in the short run. Conversely, various measures of uncertainty significantly increase in response to an identified long-run negative shock to aggregate economic activity.

Gilchrist, Yankov and Zakrajsek (2009) reach a similar conclusion from asset prices. They construct commercial credit spreads from a large panel of US nonfinancial firms, isolate the component that is orthogonal to those firms' stock prices and to contemporaneous macroeconomic conditions, and show that this component explains in a VAR framework a significant fraction of the variability of aggregate employment, but at medium horizons. A natural source of credit spread shocks are uncertainty shocks mediated by agency frictions in financial markets. Arellano, Bai and Kehoe (2011) and Di Tella (2012) argue that positive shocks to both idiosyncratic and aggregate uncertainty generate, through borrowing constraints, recessions of plausible magnitude.

Finally, a few authors have followed our early lead and entertained the hypothesis that changes in cross-sectional dispersion may be the result, and not the cause, of business cycles. Cui (2012) explains increasing dispersion in firm-level TFP in recessions through the combination of credit frictions, which prevent firms in good shape from buying assets, with the desire of distressed firms to deleverage and unload costly debt before selling their assets and used

strongly correlates with unemployment. Its explanatory power for unemployment, however, is quantitatively weak. Davis, Haltiwanger and Schuh (1996, Figure 5.5) find that the dispersion of employment growth rates across manufacturing establishments is significantly larger in 1982 (recession) than in 1978 (expansion).

capital. The slowdown in used capital reallocation, indeed observed in recessions, expands TFP dispersion. Closer to our work is the hypothesis in D’Erasmus and Moscoso-Boedo (2012). When a negative aggregate TFP shocks hits, firms find it optimal to scale down intangible investment in market access, such as marketing, advertising, brand development, or organization development. This drop in market investment reduces the spectrum and size of the markets that the firm can access, which provides less averaging and exposes the firm to more risk. The motivation is their empirical finding in a firm-level dataset that firm-level volatility in sales is countercyclical while intangible investment is procyclical, and more generally idiosyncratic volatility and intangible investment are negatively correlated across firms. As in our work, they emphasize demand shocks at the firm level and first-moment aggregate shocks. In our model, the “investment” in market access is price experimentation itself.

3 The economy

Time is discrete. There exist a continuum of differentiated varieties of a perishable consumption good or service, indexed by (j, k) , $j \in (J_1 \cup J_2)$, $k \in [0, 1]$. Each variety is produced by a single firm using only labor: h units of labor time produce $z(h - \bar{h})$ physical units, where $\bar{h} \geq 0$ is minimum employment required to keep the firm in operation, and z is aggregate productivity, which follows a Markov process.

Firms are owned by fully diversified, identical individuals, whose preferences for consumption C and work time H are expressed by the utility function

$$\log C - \chi \frac{H^{1+\eta}}{1+\eta}$$

where

$$C = \sqrt{C_1 C_2}$$

is a composite consumption good, and for $i = 1, 2$

$$C_i = \left(\int_{J_i} \int_0^1 \alpha_k^{\frac{1}{\sigma_i}} c_{j,k}^{\frac{\sigma_i-1}{\sigma_i}} dk dj \right)^{\frac{\sigma_i}{\sigma_i-1}} \quad (1)$$

is a CES aggregator of quantities $c_{j,k} \geq 0$ consumed of varieties $(j, k) \in J_i \times [0, 1]$. Here $\alpha_k \geq 0$ are preference weights. Of the two indices (j, k) that identify a variety, j is time-invariant, while k , which determines the weight α_k , is drawn every period independently over time and of j from a uniform distribution on $[0, 1]$. The draw k is the same for all consumers. Thus, the weight α_k acts as variety-specific preference shock. We identify varieties with a bi-variate

index j, k in order to allow for preference shocks (k) independent of the identity and price of the variety (j).

There are two types, or baskets, of varieties: “specialties” $j \in J_1$, hard to substitute for each other, with elasticity of substitution across varieties $\sigma_1 > 1$, and “generics” $j \in J_2$, mutually substitutable, with higher elasticity of substitution $\sigma_2 > \sigma_1$. We denote by γ_i the measure of type i varieties, the Lebesgue measure of set J_i . Finally, $\eta > 0$ is the inverse Frisch elasticity of labor supply and $\chi > 0$ the relative disutility of labor.

Each consumer supplies labor H on a competitive market at wage rate w , and receives profits Π from the firms she owns. She chooses every period consumption of all varieties j, k , sold at prices $P_{j,k}$, and labor supply, to maximize utility subject to her budget constraint. This is a static problem solved anew each period, because consumers have no access to capital markets in this economy.

Each period, each firm (j, k) posts a price $P_{j,k}$ for its variety and hires labor to produce and serve all demand at that price. The firm’s objective is to maximize the expected present value of flow profits, net of labor costs, including the fixed “operation” cost $\Psi = w\bar{h}$ per period, discounted with factor $\beta \in (0, 1)$.⁵ The discount factor incorporates both time preferences of the households, who own the firms, and a (constant) chance of exogenous exit of the firm due to some idiosyncratic destruction shock. When the firm (variety) exits, its continuation value is normalized to zero.

The novel element in our model is the information structure. Each period, the representative consumer knows her preferences, so she draws of all k ’s within each basket i , before supplying labor and making purchases; so, she knows how easily substitutable in her own preferences each variety j is. Conversely, the firm does not observe its own identity (j, k), thus whether it is a high- or low-market power firm (is $j \in J_1$ or $\in J_2$?) and the preference weight α_k that the consumer puts on its variety. For example, the consumption good is fruit, the units of all varieties are pounds of fruit, firm j produces kiwis, but does not know whether consumers see kiwis as similar to such staples as apples and oranges, difficult to substitute, or as another kind of berries, which are more price sensitive. And the consumer likes some type of kiwis (j, k) better than others (j, k'), randomly every period.

When a firm enters the industry, Nature draws the basket i to which the variety produced by the entering firm belongs to be $i = 1$ with chance λ_0 , which is also the firm’s prior belief that it is of type 1. The price quoted affects sales through the elasticity of demand, but preference shocks create demand level shocks, which hide the true elasticity. Over time, the firm observes

⁵We are implicitly assuming that firms, unlike households, have access to a perfect capital markets, so that only the expected present value, and not the timing, of profits matters. A firm can operate as long as the expected continuation value more than compensates for any current loss.

the prices it quotes and the resulting volumes of sales, and learns from them the price elasticity of the demand it faces.

We choose to make the elasticity σ_i and not the scale α of the demand curve the persistent, unobserved component that firms are interested in estimating. The reason is that, in the CES preference setup, the scale of demand does not affect the optimal price, and experimentation is “passive”. In contrast, when demand elasticity is unknown, the firm can trace the demand curve, up to noise, by changing the price. Thus, the firm faces an interesting trade-off between current profits and information about the optimal markup to charge in the long run.

4 Consumer optimization

The complete description of the consumer-worker problem is:

$$\begin{aligned} \max_{H, \{c_{j,k}\}_{(j,k) \in A}} \quad & \frac{1}{2} \sum_{i=1}^2 \frac{\sigma_i}{\sigma_i - 1} \log \left(\int_{J_i} \int_0^1 \alpha_k^{\frac{1}{\sigma_i}} c_{j,k}^{\frac{\sigma_i-1}{\sigma_i}} dk dj \right) - \chi \frac{H^{1+\eta}}{1+\eta} \\ \text{s.t.} \quad & \sum_{i=1}^2 \int_{J_i} \int_0^1 P_{j,k} c_{j,k} dk dj = Y \\ & Y = wH + \Pi \end{aligned}$$

The consumer chooses labor supply H and earns total income Y , then she decides how much income to spend on each basket and how to allocate that amount within each basket across varieties.

First, we solve the “outer” problem of choosing labor and allocating spending between baskets. For ease of presentation, we guess that there exist price indices \bar{P}_i so that we can write the budget constraint as

$$\bar{P}_1 C_1 + \bar{P}_2 C_2 = wH + \Pi$$

We will later verify this guess and solve for \bar{P}_i . The problem becomes

$$\max_{C_i, H} \log \sqrt{C_1 C_2} - \chi \frac{H^{1+\eta}}{1+\eta} + \nu (wH + \Pi - \bar{P}_1 C_1 - \bar{P}_2 C_2)$$

Taking FOCs for C_i

$$\bar{P}_1 C_1 = \bar{P}_2 C_2 = \frac{1}{2\nu}.$$

Using the budget constraint, due to the Cobb-Douglas nature of the “outer” preference aggregator, the consumer spends $Y/2$ on each basket, independently of prices:

$$\bar{P}_1 C_1 = \bar{P}_2 C_2 = \frac{Y}{2} = \frac{wH + \Pi}{2}$$

Taking a FOC for labor supply and using the above equations to solve out for the multiplier ν ,

$$\chi H^\eta = \nu w = \frac{w}{Y} = \frac{w}{wH + \Pi} \quad (2)$$

we obtain an equation solved implicitly by labor supply $H^s(w, \Pi)$.

Next, the consumer optimally allocates her $Y/2$ budget among varieties (j, k) within basket i . The FOC for an optimal $c_{j,k}$ is

$$\frac{1}{2} \frac{\alpha_k^{\frac{1}{\sigma_i} - \frac{1}{\sigma_i}} c_{j,k}^{-\frac{1}{\sigma_i}}}{\int_{J_i} \int_0^1 \alpha_k^{\frac{1}{\sigma_i} - \frac{1}{\sigma_i}} c_{j,k}^{-\frac{1}{\sigma_i}} dj dk} = \nu P_{j,k}$$

So the consumer equates the ratio $\alpha_k^{\frac{1}{\sigma_i} - \frac{1}{\sigma_i}} c_{j,k}^{-\frac{1}{\sigma_i}} / P_{j,k}$ between any two j, k and j', k' in the same basket, and

$$c_{j',k'} = c_{j,k} \frac{\alpha_{k'}}{\alpha_k} \left(\frac{P_{j,k}}{P_{j',k'}} \right)^{\sigma_i}$$

Multiplying this equation through by $P_{j',k'}$ and integrating over $j', k' \in J_i$ keeping j, k fixed

$$\frac{Y}{2} = \int_{J_i} \int_0^1 c_{j',k'} P_{j',k'} dj' dk' = \frac{c_{j,k}}{\alpha_k P_{j,k}^{1-\sigma_i}} \int_{J_i} \int_0^1 \alpha_{k'} P_{j',k'}^{1-\sigma_i} dj' dk'$$

we obtain the demand for variety j, k as a function of prices and the total budget Y

$$c_{j,k|j \in J_i} = \alpha_k \frac{P_{j,k}^{-\sigma_i}}{\bar{P}_i^{1-\sigma_i}} \frac{Y}{2} \quad (3)$$

where we define the price index:

$$\bar{P}_i^{1-\sigma_i} = \int_0^1 \alpha_{k'} \left(\int_{J_i} P_{j',k'}^{1-\sigma_i} dj' \right) dk'. \quad (4)$$

Using this expression into the demand (3) for each variety, and the resulting expression into (1) for the basket C_i , and using (4) for \bar{P}_i again, after some algebra we verify the guess $\bar{P}_i C_i = Y/2$.

We invoke a Law of Large Numbers for distribution (Glivenko-Cantelli theorem) and independence of the draws of k from j to conclude that the price index \bar{P}_i is independent of the actual realization of the k draws, and every firm knows each \bar{P}_i . Intuitively, the demand shocks α_k average out, because there are a continuum of them for each variety of type j , although the firm does not observe the specific realizations of the demand shocks. There is no correlation between demand shock $\alpha_{k'}$ and price quoted $P_{j',k'}$, because the firm quotes the price before the demand shock is realized (in a i.i.d. fashion), and anyway the firm does not observe the demand

shock even ex post; it only observed realized sales. So $P_{j',k'}$ is independent of k' . Without loss in generality, we can normalize the scale of demand shocks so that $\int_0^1 \alpha_{k'} dk' = 1$, so that

$$\bar{P}_i^{1-\sigma_i} = \int_{J_i} P_{j',k'}^{1-\sigma_i} dj' dk'.$$

Note that if the measure γ_i of active firms of type i , the Lebesgue measure of set J_i , increases, given individual prices, the rescaled price index $\bar{P}_i^{1-\sigma_i}$ increases, because it is a sum of inverse prices, not a mean. Then, the demand $c_{j,k|j \in J_i}$ for each variety of type i falls, due to more intense competition for the budget spent on basket J_i when this includes more varieties, a standard preference for variety effect. In contrast, the demand for varieties of the other type is unaffected, because the shopper always splits her budget equally between the two baskets.

5 Firm optimization

5.1 Bayesian updating

A firm enters the industry with a prior belief λ_0 that its variety has low demand elasticity σ_1 (high mark-up). Each period the firm quotes a price $P_{j,k}$ and observes the quantity $Q_{j,k}$ it sells. But the firm knows neither the customer's preference shock (α_k) nor its own demand elasticity (σ_i). Taking logs in the demand function (3), a firm j, k that is really of type i and receives preference shock α_k (but does not know either i or k) observes physical sales of

$$q_{j,k} = -\sigma_i p_{j,k} + y + \mu_i + \varepsilon_k \tag{5}$$

where

$$q_{j,k} = \log Q_{j,k}, p_{j,k} = \log P_{j,k}, y = \log Y, \varepsilon_k = \log \alpha_k, \bar{p}_i = \log \bar{P}_i, \mu_i = (\sigma_i - 1) \bar{p}_i - \log 2$$

Here the demand shock ε_k hides from the firm its true price elasticity σ_i . The term μ_i summarizes aggregate variables that also determine demand, namely the log price \bar{p}_i index quoted by the relevant (same type i) competitors, which depends on the measure of varieties that consumers buy. Therefore, given a log-price quoted $p_{j,k}$, demand can be high for one of three reasons: the variety is hard to replace in consumer preferences ($j \in J_1$), the consumer places a large weight on the variety in her preferences this period (high ε_k), and/or the variety has few competitors, so \bar{p}_i (which is an inverse log price index) is high and the consumer spends on (j, k) much of her total budget $Y/2$ for the basket.

Without loss in generality, we choose a c.d.f. F and assume that preference shocks are

$$\alpha_k = e^{F^{-1}(k)}$$

Recall that $k \sim U [0, 1]$, so F is the c.d.f. of the log-demand shock ε_k

$$\Pr (\varepsilon_k \leq \varepsilon) = \Pr (\log \alpha_k \leq \varepsilon) = \Pr (F^{-1}(k) \leq \varepsilon) = \Pr (k \leq F(\varepsilon)) = F(\varepsilon).$$

The only restriction on the support of F is the normalization $\int_0^1 \alpha_k dk = 1$.

From now on we drop the variety index j, k in (5) for convenience, so y, p, q are public information, while i and ε are only observed by the consumer. A firm that currently believes to be of type 1 with probability λ , quotes price p and observes sales q and aggregate income y , all in logs, updates by Bayes rule its belief to:

$$\lambda' = B(\lambda, p, q, Y) = \left[1 + \frac{1 - \lambda F'(q - \mu_2 + \sigma_2 p - y)}{\lambda F'(q - \mu_1 + \sigma_1 p - y)} \right]^{-1} \quad (6)$$

where $F(\tilde{q} - \mu_i + \sigma_i p - y) = \Pr(q \leq \tilde{q} | i)$. The distribution of demand shocks F does not depend on aggregate income $Y = e^y$, to avoid introducing an exogenous source of cyclical uncertainty, which is the whole point of our analysis.

Dynamic decision-making requires knowing the evolution of beliefs conditional on a given state i , where true sales are $q = \mu_i - \sigma_i p + y + \varepsilon$. After some algebra, the conditional (on true state i) evolution of beliefs is independent of aggregate income Y and we can write it as a function of prior beliefs, price quoted, and demand shocks as follows

$$b_i(\lambda, p, \varepsilon) = B(\lambda, \mu_i - \sigma_i p + y + \varepsilon, q, Y) = \left[1 + \frac{1 - \lambda F'(\mu_i - \mu_2 + (\sigma_2 - \sigma_i)p + \varepsilon)}{\lambda F'(\mu_i - \mu_1 + (\sigma_1 - \sigma_i)p + \varepsilon)} \right]^{-1} \quad (7)$$

As usual, the belief $b_i(\lambda, p, \varepsilon)$ that the state is 1 when the true state is $i = 1$ (2) has a positive (resp., negative) drift, and unconditionally on the state beliefs are a martingale:

$$E_\varepsilon [b_2(\lambda, p, \varepsilon)] < \lambda = \lambda E_\varepsilon [b_1(\lambda, p, \varepsilon)] + (1 - \lambda) E_\varepsilon [b_2(\lambda, p, \varepsilon)] < E_\varepsilon [b_1(\lambda, p, \varepsilon)]$$

The speed of learning, that is how quickly posterior beliefs react to sale volume, is controlled by the price through its effect on the likelihood ratio, namely the ratio of densities F' in (7). For example, in state $i = 1$, that ratio equals $F'((\sigma_1 - 1)(\bar{p}_1 - p) - (\sigma_2 - 1)(\bar{p}_2 - p) + \varepsilon) / F'(\varepsilon)$. So a very large deviation of own price from the price index of each basket, $|p - \bar{p}_i|$, either positive or negative, provides much information. Alternatively we can write it as $F'((\sigma_1 - 1)\bar{p}_1 - (\sigma_2 - 1)\bar{p}_2 + \varepsilon \left(\frac{\sigma_2 - \sigma_1}{\varepsilon}\right)) / F'(\varepsilon)$ so the direct effect of the price on the speed of learning is proportional to the difference in possible values of elasticities, divided by the scale of the noise, a signal/noise ratio. The higher this is, the stronger the information ‘productivity’ of price experimentation.

5.2 Optimal pricing

A firm quotes a price P , that the shopper, fully informed about her own preferences, compares with the price index \bar{P}_i of the immediate substitutes. Revenues are $PQ(P)$, where Q is the

demand function in (3). Costs are wh (variable) and $w\bar{h} = \Psi$ (fixed). Using the production function and assuming the firm serves all demand, the total demand for labor equals $Q(P)/z + \bar{h}$. Putting all together, realized flow profits for firm j equal revenues net of variable costs, minus fixed costs:

$$\Pi = \alpha \frac{Y P^{-\sigma_i} \left(P - \frac{w}{z}\right)}{\bar{P}_i^{1-\sigma_i}} - \Psi = \alpha \frac{Y}{2} \left(\frac{P}{\bar{P}_i}\right)^{-\sigma} \left(\frac{P}{\bar{P}_i} - \frac{w}{z\bar{P}_i}\right) - \Psi.$$

To fix ideas, we start with two special cases. First, the firm knows its demand elasticity. Second, it learns demand elasticity from sales but is myopic ($\beta = 0$). Finally, we move to dynamics with learning.

5.2.1 Perfect information

If the firm knows to be in the basket i , hence to face demand elasticity σ_i , then it also knows the demand it faces $\alpha Y P^{-\sigma_i} / 2 \bar{P}_1^{1-\sigma_i}$ up to the possibly unknown and i.i.d. preference shock α . In this case, the firm's problem has no intertemporal dimension and consists of maximizing flow profits each period. Using $E[\alpha] = \int \alpha_k dk = 1$, the firm chooses whether to pay the fixed cost Ψ to produce at all and, if so, the price that maximizes expected profits:

$$\max \left\langle 0, \max_P \frac{Y P^{-\sigma_i} \left(P - \frac{w}{z}\right)}{2 \bar{P}_i^{1-\sigma_i}} - \Psi \right\rangle$$

The optimal price is a constant markup, independent of Y

$$P_i^{PI} = \frac{\sigma_i}{\sigma_i - 1} \frac{w}{z}$$

and the firm earns maximized expected profits

$$\Pi_i^{PI} = \max \left\langle 0, \frac{Y}{2} (\sigma_i)^{-\sigma_i} \left(\frac{w}{\bar{P}_i (\sigma_i - 1) z}\right)^{1-\sigma_i} - \Psi \right\rangle$$

Since $\sigma_1 < \sigma_2$, type 1 firms have more market power.

5.2.2 Imperfect information: static optimization

Now the firm believes to be in set 1 with probability λ but is fully myopic and does not care about learning and experimentation. Again, the problem has no intertemporal dimension, and the firm maximizes each period expected profits

$$\max \left\langle 0, \max_P \left(P - \frac{w}{z}\right) \frac{Y}{2} \left[\lambda \frac{P^{-\sigma_1}}{\bar{P}_1^{1-\sigma_1}} + (1 - \lambda) \frac{P^{-\sigma_2}}{\bar{P}_2^{1-\sigma_2}} \right] - \Psi \right\rangle \quad (8)$$

The NFOC now yields an implicit function which defines the optimal policy $P^{MY}(\lambda)$. This does not depend on the type j, k of the firm, which is unobserved by the firm when choosing the price, but only on its belief. We know that perfect information prices are ranked $P_1^{PI} > P_2^{PI}$, because $\sigma_1 < \sigma_2$. In the Appendix, we use a monotone comparative statics argument to show that $P^{MY}(\lambda)$ is monotonically increasing in the belief λ of a high mark-up: the more optimistic the firm is to offer a special variety that consumers really like, the higher the price it charges.

Proposition 1 (*Profit-maximizing price when demand elasticity is uncertain*) *The profit-maximizing price under imperfect information about the elasticity of demand, the maximizer $P^{MY}(\lambda)$ of (8), is strictly increasing, from $P^{MY}(0) = P_2^{PI}$ to $P^{MY}(1) = P_1^{PI}$, in the belief λ of a low demand elasticity.*

5.2.3 Imperfect information: dynamic optimization

The firm's Dynamic Programming problem can be written recursively with a state vector which includes an idiosyncratic state variable, belief λ , and aggregate states that we summarize in the vector ω , with Markov transition \mathcal{T} on a support Ω . The firm now maximizes the expected present value of profits discounted with factor $\beta \in (0, 1)$. Then its value is

$$V(\lambda, \omega) = \max \langle 0, W(\lambda, \omega) \rangle$$

where the first option is exit and the second is continuation:

$$\begin{aligned} W(\lambda, \omega) = & \max_P \left(P - \frac{w}{z} \right) \frac{Y}{2} \left[\lambda \frac{P^{-\sigma_1}}{\bar{P}_1^{1-\sigma_1}} + (1-\lambda) \frac{P^{-\sigma_2}}{\bar{P}_2^{1-\sigma_2}} \right] - \Psi \\ & + \beta \lambda \int_{\Omega} \left[\int V(b_1(\lambda, \log P, \varepsilon), \omega') dF(\varepsilon) \right] \mathcal{T}(d\omega', \omega) \\ & + \beta (1-\lambda) \int_{\Omega} \left[\int V(b_2(\lambda, \log P, \varepsilon), \omega') dF(\varepsilon) \right] \mathcal{T}(d\omega', \omega) \end{aligned} \quad (9)$$

To calculate beliefs b_i for each type of firm i , we use the stochastic law of motion (7) derived earlier. The resulting optimal policy is denoted by $P^*(\lambda, \omega)$.

Since the firm enjoys more market power if it is in set J_1 , we should expect V to be increasing. Then exit occurs at a belief $\underline{\lambda}_\omega \in [0, 1]$, if any, such that value matching holds

$$V(\underline{\lambda}_\omega, \omega) = 0 \quad (10)$$

and the firm continues as long as the belief remains above $\underline{\lambda}_\omega$. If no interior $\underline{\lambda}_\omega$ exists, then the firm either never exits or never enters. At $\lambda = 1$ there is perfect information and the firm marks up over the wage the usual way: $P^*(1, \omega) = P_1^{PI}$. At the other extreme, $V(0) = 0 > W(0)$

if the firm exits before being sure to have low market power, which requires a large enough operation cost Ψ to make $\Pi_2^{PI} < 0$.

The state vector of the firm ω includes exogenous aggregate states, such as labor productivity z , and four endogenous states, the price indices of the two baskets, \bar{P}_1 and \bar{P}_2 , the wage rate w , and aggregate income Y . The laws of motion of the last four are determined by general equilibrium conditions, to which we now turn.

6 General equilibrium

We now reintroduce variety index (j, k) for aggregation. Equilibrium consistency requires for each basket i

$$\bar{P}_i^{1-\sigma_i} = \int_0^1 \alpha_k \int_{J_i} P_{j,k}^{*1-\sigma_i} dj dk$$

The optimal price $P^*(\lambda, \omega)$ does not depend on the place of the variety in preferences j, k , but only on the state of the firm profit maximization problem λ, ω . Because the current belief λ depends on past sales and preference shocks, not on the current preference shock k , denoting by ϕ_i the cross-section c.d.f. of posterior beliefs of firms of type i ,

$$\bar{P}_i^{1-\sigma_i} = \int_0^1 \alpha_k \int_{J_i} [P^*(\lambda, \omega)]^{1-\sigma_i} d\phi_i(\lambda) dk = \int_{J_i} [P^*(\lambda, \omega)]^{1-\sigma_i} d\phi_i(\lambda)$$

where we use the independence of the k draws from j and $\int_0^1 \alpha_k dk = 1$.

Market-clearing for each variety is guaranteed by the fact that firms incorporate the demand function in their pricing decisions, and then produce the resulting quantity demanded by the shopper, which is $c_{j,k|j \in J_i}$. To produce this amount, the firm needs to hire $c_{j,k|j \in J_i}/z + \bar{h}$ units of labor. Market clearing in the labor market requires that labor supply $H^s(w, \Pi)$ in (2) equals labor demand $H^d(w, \Pi)$, which is the sum of labor demand by all firms

$$H^d = \sum_{i=1}^2 \int_{J_i} \left(\int_0^1 \frac{c_{j,k|j \in J_i}}{z} dk + \bar{h} \right) dj$$

Using the demand function for variety j and the budget constraint of the household $Y = wH + \Pi$,

$$\begin{aligned} H^d - \bar{h}(\gamma_1 + \gamma_2) &= \sum_{i=1}^2 \int_{J_i} \int_0^1 \alpha_k \frac{Y}{2z} \frac{P_j^{-\sigma_i}}{\bar{P}_i^{1-\sigma_i}} dk dj \\ &= \frac{Y}{2z} \sum_{i=1}^2 \frac{\int_{J_i} \int_0^1 \alpha_k [P^*(\lambda, \omega)]^{-\sigma_i} dk}{\bar{P}_i^{1-\sigma_i}} d\phi_i(\lambda) = \frac{wH^d + \Pi}{2z} \sum_{i=1}^2 \frac{\int_{J_i} [P^*(\lambda, \omega)]^{-\sigma_i} d\phi_i(\lambda)}{\int_{J_i} [P^*(\lambda, \omega)]^{1-\sigma_i} d\phi_i(\lambda)} \end{aligned}$$

We choose labor as the numeraire, $w = 1$, so this equation and labor supply pin down profits Π and employment H , thus aggregate income $Y = H + \Pi$. The state vector is $\omega = \{z, \phi_i(\cdot)\}$, potentially infinitely-dimensional.

To close equilibrium, we need to pin down the belief distribution ϕ_i and measures γ_i of active firms of each type. Because firms exit exogenously and (potentially) endogenously, we need to specify an entry structure. We assume that Nature assigns to a new firm type 1 with probability equal to the prior belief λ_0 , and provide more details below.

6.1 Perfect information

In this case, the distribution of beliefs is simple, concentrated on atoms $\phi_1(1) = \phi_2(0) = 1$. Therefore, the optimal pricing policy is the usual mark-up equation for P_i^{PI} . Since the consumer needs varieties of both types and spends half of her income on each basket, there cannot be an equilibrium where no firm of a given type operates, as a single firm could sell at an arbitrarily large price an arbitrarily small quantity (thus at negligible variable cost), obtain revenues $Y/2$ and make positive expected profits as long as $Y/2 > \Psi$. With perfect information, in those circumstances equilibrium requires that exit of varieties occurs until $\Pi_i^{PI} = 0$.

Firms of the same type quote the same price, so the price index equals this common price

$$\bar{P}_i^{PI} = P_i^{PI} = \frac{\sigma_i}{\sigma_i - 1} \frac{w}{z} \gamma_i^{\frac{1}{1-\sigma_i}}$$

Replacing this price in the flow profits function, and taking expectations w.r. to the demand shock, we can solve for expected profits as a function of variables that the firm takes as given:

$$\Pi_{j \in J_i}^{PI} = \frac{Y}{2\sigma_i \gamma_i} - \Psi.$$

This expression is declining in the measure γ_i of varieties in basket i : the more there are, the stronger competition for the share $Y/2$ of income going to the basket, the lower profits for each variety.

Aggregating across all varieties in both baskets and averaging out demand shocks, total profits rebated to the consumer are

$$\Pi = \frac{Y}{2} \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} \right) - \Psi(\gamma_1 + \gamma_2).$$

If the proportion $\gamma_2/(\gamma_1 + \gamma_2)$ of type 2 firms active in the market is sufficiently low relative to price elasticities $\sigma_2/\sigma_1 > 1$, a type 2 firm will earn higher profits than a type 1 firm, despite the lower markup, due to the larger scale of demand per firm. With free entry, the measure γ_i of firms of each type i cannot exceed an upper bound which drives $\Pi_{j \in J_i}^{PI}$ to zero in each state.

6.2 Imperfect information: steady state

With z and all other aggregate variables constant, the stationary distributions ϕ_i will be such that the flow of firms of type 1 that exit is a fraction λ_0 of all exiting firms, and entry and exit balance for all firms, so they balance for each type of firms. Let

$$n_i = \int_0^1 \Pr(\varepsilon : b_i(\lambda, p, \varepsilon) \leq \underline{\lambda}) d\phi_i(\lambda)$$

be the endogenous probability of exit of firms of type i , and δ the exogenous exit probability of all firms. Since type 2 firms are less profitable, they will exit faster on average (b_2 is a supermartingale and b_1 a submartingale). The balanced flows equation

$$\frac{\gamma_1(\delta + n_1)}{\gamma_1(\delta + n_1) + (1 - \gamma_1)(\delta + n_2)} = \lambda_0$$

equates the fraction of exiting firms of type 1 to the fraction of entering firms of type 1. We can solve this equation for the steady state share of type 1 firms γ_1 given exit probabilities,

$$\gamma_1 = \left(1 + \frac{\delta + n_1}{\delta + n_2} \frac{1 - \lambda_0}{\lambda_0}\right)^{-1}$$

and then the actual flow of firms that exits each period is

$$\gamma_1(\delta + n_1) + (1 - \gamma_1)(\delta + n_2) = \frac{(\delta + n_1)(\delta + n_2)}{\lambda_0(\delta + n_2) + (\delta + n_1)(1 - \lambda_0)}$$

6.3 Imperfect information: stochastic equilibrium

When the economy is hit by aggregate shocks to z , the distributions ϕ_i of beliefs are state variables for each firm. To avoid carrying the total measure of active firms as an additional state variable, we think of the following organization of the industry. There is a fixed unit measure of “slots” (location, land lot, shelf space in a store), that products can occupy. Each firm produces one variety and can occupy one slot. The operating cost Ψ is the rental cost of the slot. When she goes shopping, the consumer can only find a maximum measure of slots/products, that we normalize to one. The variety index j is also the slot index. Firms of type 1 take the lower-numbered slots, thus their set is $J_1 = [0, \gamma_1]$ and firms of type 2 take the high-numbered slots, so their measure is $\gamma_2 = 1 - \gamma_1$, and their set $J_2 = (\gamma_1, 1]$. After firms exit, for exogenous or endogenous reasons, new firms instantaneously enter, until slots are saturated. Employment H , income Y , and the proportions of operating firms $\gamma_1 = 1 - \gamma_2$ are all functions of the aggregate state ω .

7 Numerical illustrations

In our leading parameterization, the log-demand shock ε is normal $N(m, s^2)$, so

$$B(\lambda, p, q, Y) = \left[1 + \frac{1-\lambda}{\lambda} \exp \left\{ \frac{(q - \mu_1 + \sigma_1 p + y - m)^2 - (q - \mu_2 + \sigma_2 p + y - m)^2}{2s^2} \right\} \right]^{-1}.$$

After some algebra in Appendix, denoting an independent noise term by $\zeta \sim N(0, s^2)$:

$$b_i(\lambda, p, \zeta) = \left[1 + \frac{1-\lambda}{\lambda} \exp \left\{ \left(-\frac{1}{2} \right)^{\mathbb{I}(i=2)} \left[\left(\frac{\zeta}{s} \right)^2 - \left(p \frac{\Delta\sigma}{s} - \frac{\Delta\mu}{s} + \frac{\zeta}{s} \right)^2 \right] \right\} \right]^{-1}$$

where \mathbb{I} is the indicator function. So the true state i just determines the sign of the log-likelihood ratio, which is the term in curly brackets. The speed of learning, that is how quickly posterior beliefs react to sale volume, is controlled by the price through the signal/noise ratio $\Delta\sigma/s$: the farther apart the two possible values of the elasticity, relative to the noise in demand, the more informative are sales. Note also that

$$p \frac{\Delta\sigma}{s} - \frac{\Delta\mu}{s} = \frac{(p - \bar{p}_1)(\sigma_1 - 1) - (p - \bar{p}_2)(\sigma_2 - 1)}{s}$$

so a very large $|p - \bar{p}_i|$, either positive or negative, provides much information. In order to learn, firms have to price differently than potential competitors in either basket, and the more pronounced the difference in elasticities the more pronounced the difference in sales for any amount of noise ε .

7.1 The incentives to price-experiment

The firm trades off the desire to price at the static (“myopic”) optimal mark-up P^{MY} given current beliefs, against the goal of making posterior beliefs move and learning the truth, to price closer to the perfect information optimal markup P^{PI} in subsequent periods. Note that P^{MY} depends on what other firms choose (through \bar{P}_1, \bar{P}_2), because the firm is unsure against whom it is competing, while P^{PI} is the usual mark-up $\sigma/(\sigma - 1)$ over marginal cost, independent of competitors’ prices.

The objective function of the dynamic experimentation problem (9) is the sum of two terms, a static profit function and a continuation value. The former is concave in the price (in deviation from the average price index), and peaks at an intermediate value P^{MY} of the price. The latter is generally convex and U-shaped in the same price (deviation), because either a very high or a very low price is very informative and raises the future value, while a small deviation from others’ prices teaches the firm almost nothing, and makes the continuation less valuable.

The firm trades off these two incentives, a middle price that maximizes static profits and an extreme price that maximizes the value of information. The sum of the two terms on the RHS of (9), a concave function plus a convex function, is generally a double-peaked function, with two local maxima, a low price and a high price, with the myopic price P^{MY} somewhere in between.⁶ The twin peaks introduce a potential discontinuity in the optimal policy function. For reasons explained below, in our model the high price is the one that the firm always selects, so experimentation introduces an upward bias in prices and the optimal policy is continuous. But it is conceivable that in other demand systems, or different parameterizations or calibrations of the problem, the firm may optimally alternate between the low price and the high price, depending on the values of the state variables.

Standard in Bayesian learning problems, the value function V of the optimal pricing-experimentation problem is convex. Beliefs are a martingale, and more information cannot hurt, as it can always be ignored. So, for any price:

$$\begin{aligned} V(E_\varepsilon[\lambda b_1(\lambda, \log P, \varepsilon) + (1 - \lambda) b_2(\lambda, \log P, \varepsilon)], \omega) &= V(\lambda, \omega) \\ &\leq E_\varepsilon[\lambda V(b_1(\lambda, \log P, \varepsilon), \omega) + (1 - \lambda) V(b_2(\lambda, \log P, \varepsilon), \omega)] \end{aligned}$$

The firm benefits from a mean preserving spread in posterior beliefs, that it can generate by deviating from competitor's prices, as sales are then very informative. The identity of the relevant competitors (the type of the firm) is unknown, but the firm has beliefs about it.

We expect the gains from experimentation to increase with:

- future income Y' , because flow profits are scaled by income and information is valuable in the future
- current income Y if its process is persistent, as a higher Y signals a higher future Y' ;
- the signal/noise ratio $\Delta\sigma/s$, as experimentation yields more “information bang for the buck”;
- the operation cost Ψ , because information helps the firm to exit at the appropriate time and to stop paying Ψ when not worth it.

The cost of experimentation is an expected loss in current profits. We expect this loss to increase with:

- current income Y , which scales current profits, hence the cost of making pricing “mistakes” in the short run;

⁶We thank Veronica Guerrieri for pointing out this property of the problem to us.

- the level of the elasticities σ_i , given their difference $\Delta\sigma$, as a given price deviation to attain a certain amount of learning has a more dramatic impact on profits the more elastic is demand;
- the operation cost Ψ , as the firm pays it before receiving orders and producing, thus is the fixed cost of not leaving the market to trying and learn.

Note that the current state of demand Y affects both the gains and the costs of experimentation. So experimentation increases in expected output growth Y'/Y : the opportunity cost of poor sales today is low relative to future potential sales. Temporarily bad times are a motive to experiment, even without operation costs Ψ and exit. The role of Ψ is to make experimentation more valuable for firms in bad shape, especially in bad aggregate economic conditions, so that the effect of fundamentals on the incentives to experiment work both for aggregate and idiosyncratic shocks.

More complex are the effects of the belief λ of high market power on the firm's incentives to experiment. When λ is close to 1, the firm is happy to continue in the market and to sell at a price approximately equal to the static markup $P_1^{PI} = (w/z)\sigma_1/(\sigma_1 - 1)$. Unexpected sales are attributed almost entirely to bad luck, i.i.d. spending shocks α_k , so λ moves very slowly, and the price with it. As λ approaches $1/2$ from above, the firm is increasingly uncertain and gets the most bang for the buck in terms of learning from any unexpected sales, so the incentives to experiment increase. As λ falls further and approaches the cutoff $\underline{\lambda}$, if this is much below $1/2$, the looming possibility of exit generates two conflicting effects on the firm's incentives to experiment. Limited liability and exit make the firm's value function locally more convex in beliefs: the firm "gambles for resurrection," because if the results of the price experiments are disappointing the firm can always exit and avoid the consequences. On the other hand, as the chance of continuation in the market declines, the likelihood of being able to use fresh information in the future declines, the firm discounts the future more heavily, and minds static profits more. Because the aggregate state of the economy z directly influences the exit cutoff $\underline{\lambda}$, hence the chance of exit for any given λ , it has similar contrasting effects on experimentation, in addition to those illustrated earlier.

To assess these complex and conflicting effects, we resort to numerical examples. First, we explore a firm's optimal pricing behavior when facing stochastically evolving average output Y and price indices \bar{P}_i , without requiring consistency between these aggregates and actual pricing behavior. This is a partial equilibrium exercise, to understand a firm's intertemporal incentives when facing shocks to average demand and to competitors' prices. Second, we study steady state equilibrium where aggregate productivity z is constant, entry and exit keep the measure

of active firms in balance, and price indices are \bar{P}_i consistent with average prices by active firms, so all markets clear. We are still working on full stochastic equilibrium.

7.2 Partial equilibrium

The model period is one quarter. We choose the following parameter values.

β	σ_1	σ_2	s^2
0.98	8	10	0.1

The discount factor β includes a 1% exogenous exit rate. For aggregate dynamics we normalize mean output to 1 and run a partial equilibrium exercise. We let price indices rise and fall with aggregate output. We choose $w = 1$, labor as the numeraire, and assume a change in z around $z = 1$ such that:

real output Y	Prob. Y switches	oper. cost Ψ
0.95	0.1	0.058
1.05	0.1	0.058

This is a partial equilibrium exercise because we allow price indices to move up and down by 0.5%, inversely with output. The average \bar{P}_i across states equals the average perfect information mark-up price for firms of type i , namely $\bar{P}_1^{PI} = 8/7$ and $\bar{P}_2^{PI} = 10/9$. Then we solve for the optimal pricing policy, and simulate a panel of firms to verify how the dispersion of its price changes comoves with the aggregate output.

Figure 1 shows the value function $V(\lambda, Y)$ of the firm's problem as a function of the belief λ that the variety has low elasticity σ_1 , when the current state of aggregate economic activity Y takes each of its two possible values. As mentioned, the value is convex, the upper envelope of a smooth increasing and convex function W and the zero outside option when choosing exit, which occurs where the $W = 0$.

As mentioned, the firm can learn by choosing a large deviation of its own price from the rescaled price indices \bar{p}_i . Given the structure of preferences, under perfect information a firm of type 1 only cares about \bar{p}_1 (and vice versa): only \bar{p}_1 is relevant to the firm's decisions, any change in the other price index \bar{p}_2 will only reduce proportionally the real income spent on basket 2, leaving spending on each basket unchanged. With imperfect information, the firm knows both values of \bar{p}_1 and \bar{p}_2 in equilibrium, but not which one is relevant to its sales. Therefore, it can learn by quoting a price that is very high or very low relative to *either* price index \bar{p}_i , unless \bar{p}_1 and \bar{p}_2 are very different. To a first approximation, the effect of a drastic price change on learning only depends on its size and not on its sign: note that the leading term in (7) is $(p\Delta\sigma/s)^2$ for $|p|$ large. A very high price ($p > 0$), though, enjoys a unique advantage: it reduces sales, thus production costs. A low price ($p < 0$) must raise revenues much faster

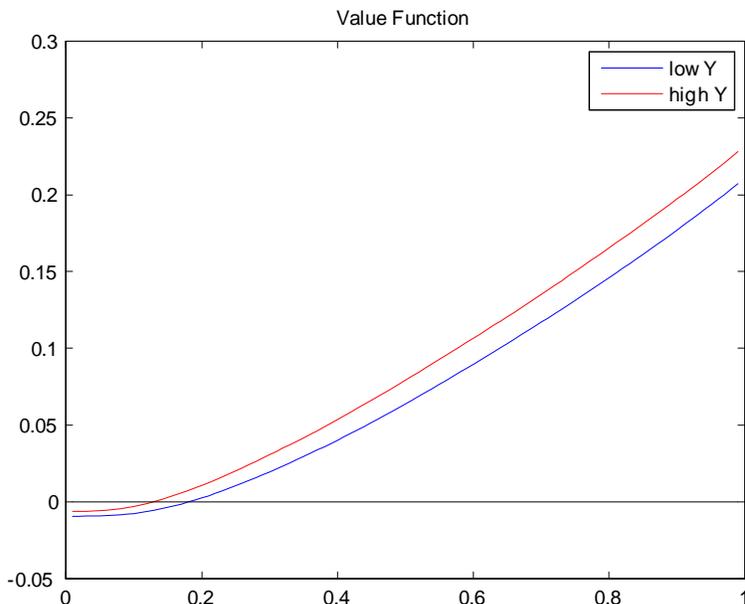


Figure 1: Value function of the firm in aggregate stochastic environment

than a higher price reduces them, in order to offset the increased costs. Whether this happens or not depends on the demand system. With isoelastic demand, this never happens. In our numerical example, we find that experimentation always takes the form of a price exceeding the myopic optimum under uncertainty $P^{MY}(\lambda)$. Imperfect information introduces a positive bias in prices. This upward pressure on prices is stronger the more pronounced experimentation. In our view, and in our numerical example, experimentation is more common in bad economic times, when spending fall, for reasons explained above.

Figure 2 illustrates the optimal policy function (price), again as a function of the belief λ that the variety has low elasticity σ_1 , when the current state of aggregate economic activity Y takes each of its two possible values. The myopic pricing function P^{MY} (not shown) is increasing. In contrast, when approaching exit, the firm experiments by pricing high, to “read” demand elasticity while saving on production costs. The likelihood of exit makes the continuation value locally more convex, raising the returns to experimentation. When exit becomes almost sure, then the likelihood of being able to use information in the future declines, and the price falls back quickly towards the myopic level. These conflicting incentives generate the “hump” in the pricing function visible in Figure 2.

Since the operations cost Ψ is inelastic to changes in aggregate demand Y , it is more onerous, and exit is more likely, when Y is low. Also, profits are scaled by Y , so a low Y makes the

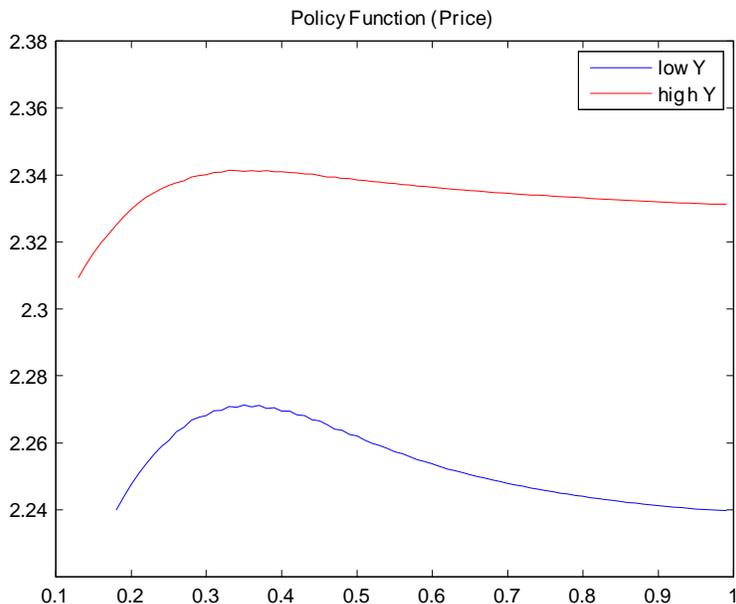


Figure 2: Pricing policy of the firm in aggregate stochastic environment

profit loss from extreme pricing less severe. Correspondingly, the hump is more pronounced in bad times. When Y is high, firms stick to a fairly uniform price, so that learning requires deviating from both, experimenting a bit and then cutting prices as evidence that the product does not sell well accumulates, and the belief λ falls. Since type 1 firms tend to have beliefs closer to 1, and type 2 firms closer to the exit cutoff, the prices of the two baskets \bar{P}_1 and \bar{P}_2 differ especially when Y is low. Possibly, the inferior varieties of type 2 may have a higher average price, as firms experiment. Thus, it is easier to learn by changing prices around, and the benefits from doing so justify the cost in terms of forgone flow profits.

When Y switches, there is a jump in all prices, caused by the discreteness of the aggregate state. The effect, though, is not symmetric. When the economy expands, firms that were closer to the exit cutoff cut their prices by more, as the incentives to experiment diminish. When the economy slump, those same firms (low beliefs) let their prices rise more, to observe how resilient sale volumes are, then further adjust the price depending on what they observe. Prices at the low end of the belief distribution tend to decline on average, as beliefs are a martingale and the pricing function is concave, hence prices are a supermartingale.

In order to simulate this economy, we assume that upon endogenous exit by each firm at $\underline{\lambda}$ a new firm enters the market, and is assigned to basket 1 with chance $\lambda_0 = 0.5$. We simulate 50,000 firms over 500 model periods. We then calculate in each period the cross-sectional

standard deviation of price changes. Price changes, and particularly price cuts sliding “off the hump” on either side of the peak, are more frequent in bad economic times. We find that this measure of price change dispersion has a correlation of about -0.38 with aggregate economic activity Y . Berger and Vavra (2010) report correlations around -0.4 between the interquartile range of the monthly changes in CPI components and detrended industrial production. In our model, revenues and output are proportional to prices, so a standard deviation in % price changes translates directly into one of sales and output changes.

7.3 Steady state general equilibrium

We now assume that aggregate labor productivity is constant at $z = 1$. We now have to provide a full calibration of the economy, because we longer treat output and price indices as exogenous state variables. Let δ denote the chance of exogenous exit, which can be absorbed into the discount factor to solve the firm’s experimentation problem in partial equilibrium, but must be set separately in general equilibrium to compute the outflow of firms.

β	δ	σ_1	σ_2	η	λ_0	$Var(e^\varepsilon)$	Ψ	H
.99	.02	8	10	1	.4	.1	.0367	.33

We require that aggregate employment equals $1/3$ (of available time), and iterate over the cost of effort parameter χ . Specifically, we proceed as follows:

1. guess values for χ , \bar{P}_1 and \bar{P}_2 ; calculate the resulting value of aggregate output Y and employment H from market-clearing in the output market and from the labor supply equation (2);
2. solve the firm problem in steady state, and find the optimal pricing policy;
3. simulate the history of 50,000 firms, starting from a degenerate distribution of beliefs at λ_0 ; at each step, replace each firm that exits either exogenously or endogenously with a new firm; assign to this new firm type σ_1 with chance λ_0 ;
4. stop the simulation when the distribution of beliefs settles in both measure of firms active for each type and distribution of beliefs by type; compute aggregate employment and price indices \bar{P}_i adding across firms the simulated prices and resulting employment from the last round of the simulation;
5. if price indices equal the guess from step 1 and employment H equals 0.33, stop; otherwise, update χ (in the direction of $H - 0.33$) and go back to step 1 with the new price indices.

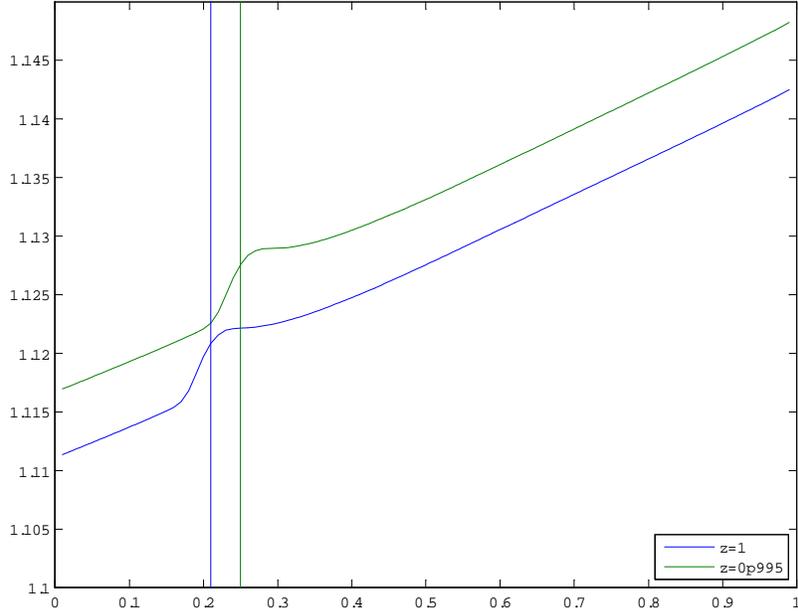


Figure 3: Steady state optimal pricing policy for two values of aggregate TFP

After finding equilibrium for $z = 1$, we fix the resulting value of work disutility χ and repeat the exercise for $z = 0.995$, this time iterating over employment H rather than preference parameter χ , until H converges to a steady state value (below 0.33).

Figure 3 illustrates the pricing policy and exit cutoff for each aggregate state. Because the shocks are to aggregate labor productivity, good prices in units of labor and exit cutoff are both countercyclical. The hump near the exit cutoff, albeit much less prominent than in the partial equilibrium exercise, is still visible, and we verified numerically that it is more pronounced in the low aggregate state. For larger beliefs, the quoted price is very close to their myopic optimal value. Barring the cyclical effects described earlier, in steady state the only reason to experiment is the desire to price according to the true demand elasticity, including the option of exit should demand turn out to be elastic.

7.4 Hazard rate of price changes

To conclude this section, we study another testable prediction of the model, the hazard rate of price changes. In both CPI and scanner data, it is well-known that prices are constant for months at a time, with the median price duration exceeding six months, depending on the treatment of sales, and the hazard rate of a price change is declining with price duration (Nakamura and Steinsson 2008), even, crucially, when eliminating cross-item heterogeneity in

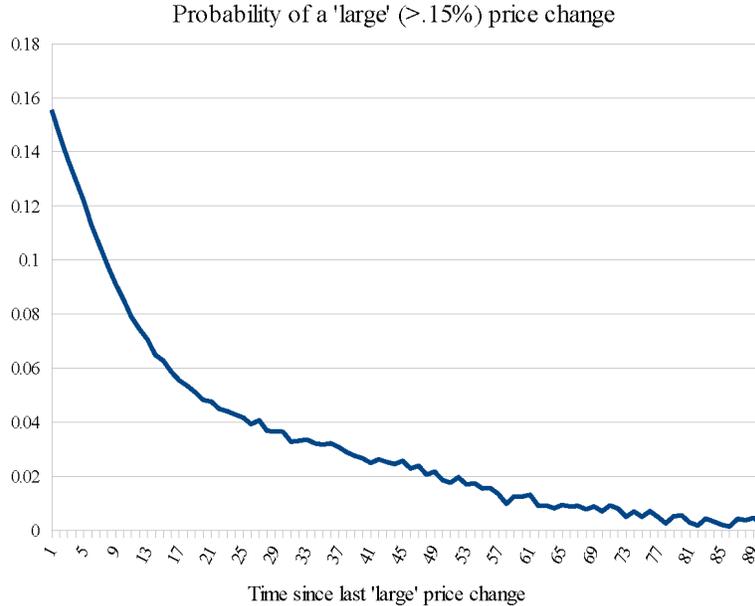


Figure 4: Hazard rate of price changes

the frequency of price adjustments and exploiting repeated spells for the same item (Vavra 2010). Our model cannot produce price stickiness, because it features no fixed costs of price adjustment. Introducing such a cost and making the optimal pricing policy a belief-dependent S,s rule would be prohibitively complex. As an approximation, we focus on ‘large’ percentage price changes, that we define to be those that exceed (in absolute value) 0.15%. In the presence of small menu costs, these would be the only price changes to be observed. In our steady state equilibrium with $z = 1$, which is calculated through simulation of 50,000 firms over 200 periods, the 0.15% threshold implies that a large price change occurs roughly every 8 months, in line with the empirical evidence on CPI price duration. We study the hazard rate of such large changes, namely the probability it occurs at duration t among all prices that have not reset until t .

Figure 4 shows that, as expected, the hazard rate is strongly declining in duration, consistently with the empirical evidence. In contrast, the Calvo model of price adjustment assumes a constant hazard rate, and menu cost models predict an increasing hazard. The information cost of Reis (2006) predicts a constant and then occasionally spiking hazard rate, while the hybrid menu cost/information cost model of Alvarez et al. (2011) generates a non-monotonic hazard rate, increasing, then spiking and decreasing. We see this as a major strength of our model of price-setting.

7.5 Stochastic general equilibrium

[Under construction.]

8 A monetary model with menu costs and price experimentation

So far we have treated labor as the numeraire, so all prices we refer to are relative prices. The attention of the literature, and the empirical evidence we presented, pertain, however, to nominal prices. As mentioned, our model cannot generate persistent exact price stickiness. To amend both, we show how to introduce money and menu costs in our model. The goal is to generate a declining hazard rate of genuine price changes. Intuitively, a firm is reluctant to change its price continuously, because of adjustment (menu) costs. When the firm does change the price, it can set it optimally so as to maximize profits but also to acquire valuable information about the demand curve it faces. Typically, the first sales observed after the price change will contain an element of surprise. As the firm refines its estimate of demand elasticity, the likelihood rises of a further price change, to take advantage of the new information. Conversely, when a price has been constant for a long time, it is likely that the firm knows its demand elasticity well, and does not need to change the price further. Hence, the declining hazard rate of price changes with duration that we observe in the data is the result of selection by unobservables (beliefs about the demand curve). The presence of menu costs, however, makes price adjustment infrequent and optimal, so especially unlikely right after an adjustment.

Preferences and information structure are as before. Money is the numeraire. P is the nominal price of the variety produced by the firm, \bar{P}_i is the price index of basket $i = 1, 2$. Money supply follows

$$\log M' = \pi + \rho \log M + \xi$$

where ξ is an i.i.d. shock.

At each point in time, the consumer optimally allocates half of expenditure to each basket $i = 1$, so

$$\bar{P}_1 C_1 = \bar{P}_2 C_2 = \frac{M}{2}.$$

Let the ideal price index

$$P^* = \sqrt{\bar{P}_1 \bar{P}_2}.$$

which equates real spending M/P^* to indirect utility. For simplicity, we fix the *real* wage to w in utility units through a perfectly elastic labor supply. We also dispose of the fixed operation

cost Ψ , and introduce a menu cost κ , in labor units, a fixed cost that must be paid to change the current nominal price.

To make the firm DP problem stationary, we define detrended money supply and prices

$$m := \frac{M}{M_0(1+\pi)^t}; \quad p := \frac{P}{M_0(1+\pi)^t}; \quad \bar{p}_i := \frac{\bar{P}_i}{M_0(1+\pi)^t}$$

Note that these are not real variables, but nominal variables in deviations from a nominal trend. It is easy to show that Bayes rule is identical to the stationary case with detrended prices (updating only depends on relative prices, obviously). Using these definitions and equalities, we express flow profits in real terms

$$\Pi(\lambda, P, \bar{P}_1, \bar{P}_2, M) = \frac{M}{2\sqrt{\bar{P}_1\bar{P}_2}} \left[\lambda \frac{P^{-\sigma_1}}{\bar{P}_1^{1-\sigma_1}} + (1-\lambda) \frac{P^{-\sigma_2}}{\bar{P}_2^{1-\sigma_2}} \right] \left(P - w\sqrt{\bar{P}_1\bar{P}_2} \right)$$

as a function of nominal detrended variables:

$$= \frac{m}{2} \left[\lambda \frac{1}{\bar{p}_1} \left(\frac{p}{\bar{p}_1} \right)^{-\sigma_1} + (1-\lambda) \frac{1}{\bar{p}_2} \left(\frac{p}{\bar{p}_2} \right)^{-\sigma_2} \right] \left(\frac{p}{\sqrt{\bar{p}_1\bar{p}_2}} - w \right)$$

We now turn to the firm's DP problem. The timing is as follows. The firm enters the period with previously quoted price P , belief λ , and last period's prices of the two baskets \bar{P}_i and money supply M . Then a new draw of money supply M' is observed by all firms, who then choose whether to change the price or not, P possibly changes to P' and \bar{P}_i to \bar{P}'_i , each firm observes its own sales, collects profits, updates beliefs. So today's profits and updating depend on M' , P' and \bar{P}'_i .

Detrended money supply follows

$$\log m' = \rho \log m + \xi$$

Given the state $(\lambda, p, \bar{p}_1, \bar{p}_2, m)$ in detrended terms, let $V(\lambda, p, \bar{p}_1, \bar{p}_2, m)$ be the value function of the firm at the beginning of the period, before the new money supply m' is drawn, firms adjust prices and then update beliefs. Let $W(\lambda, p', \bar{p}'_1, \bar{p}'_2, m')$ be the value of the firm after firms observe m' , quote detrended prices p' and \bar{p}'_1, \bar{p}'_2 , but before they observe sales and update beliefs. Remember that Π is the *expected* profit function, given current belief λ . So W is the value of the firm after new prices are quoted and before customers react. Note that, after updating λ to the new belief $\lambda' = B(\lambda, p', \bar{p}'_1, \bar{p}'_2, \zeta)$ through bayes rule B , the vector $(\lambda', p', \bar{p}'_1, \bar{p}'_2, m')$ becomes the state at the beginning of next period. Then

$$W(\lambda, p', \bar{p}'_1, \bar{p}'_2, m') = \Pi(\lambda, p', \bar{p}'_1, \bar{p}'_2, m')$$

$$+\beta E_{\zeta} [\lambda V (b_1 (\lambda, p', \bar{p}'_1, \bar{p}'_2, \zeta), p', \bar{p}'_1, \bar{p}'_2, m') + (1 - \lambda) V (b_2 (\lambda, p', \bar{p}'_1, \bar{p}'_2, \zeta), p', \bar{p}'_1, \bar{p}'_2, m')]$$

and the Bellman equation is

$$V (\lambda, p, \bar{p}_1, \bar{p}_2, m) = E_{m'} \left[\max \left\langle W \left(\lambda, \frac{p}{1 + \pi}, \bar{p}'_1, \bar{p}'_2, m' \right), \max_{p'} W (\lambda, p', \bar{p}'_1, \bar{p}'_2, m') - \kappa \right\rangle \right]$$

If the firm does not pay the menu cost κ to adjust the nominal, undetrended price P , the detrended price p declines by an amount equal to the trend inflation rate π .

8.1 Steady state

We compute general equilibrium when money growth is constant at π and experiences no shocks. The algorithm proceeds by guessing values for the price indexes of the two baskets \bar{P}_1 and \bar{P}_2 , solves the firm's DP problem, simulates a large panel of firms, including entry and exit, and computes the implied average prices of firms in each basket, which provide an update for \bar{P}_1 and \bar{P}_2 . Iterating until converges yields the desired equilibrium.

The firm's DP problem generates an optimal sS policy in two dimensions: old price P and belief λ about demand elasticity. The firm is more willing to change its price the farthest it is from the optimal price, given current beliefs, and the less extreme beliefs are, because then price changes yield more valuable information.

We now present, by way of illustration, results from a fully computed stationary general equilibrium, where price indexes have converged to initial guesses. For simplicity, we eliminate endogenous exit and allow only for exogenous exit with probability δ , and we make labor supply infinitely elastic, fixing the real wage at a normalized value of 1. The rest of the parameter values are:

β	δ	σ_1	σ_2	η	λ_0	$Var(e^\varepsilon)$	Ψ	π	κ
.99	.02	8	10	∞	.4	.1	0	.02	.01

Figure 5 shows the optimal band policy. If beliefs are either 0 or 100 (%) that the state is 1, then we are back to perfect information and obtain a standard sS policy. The optimal price increases with the belief because the mark-up is higher, so bands are skewed to the right. The width of the bands is higher for intermediate beliefs, because there news from realized sales matter more, beliefs move faster, so the larger variance of belief changes induces a standard option value of waiting effect.

When we calculate the hazard rate of price adjustment, for this parameter configuration we find in Figure 6 that the menu cost/sS force still dominates: the hazard rate rises, spikes at the duration (around 21) when inflation and shocks push the nominal price against the band, and declines after that.

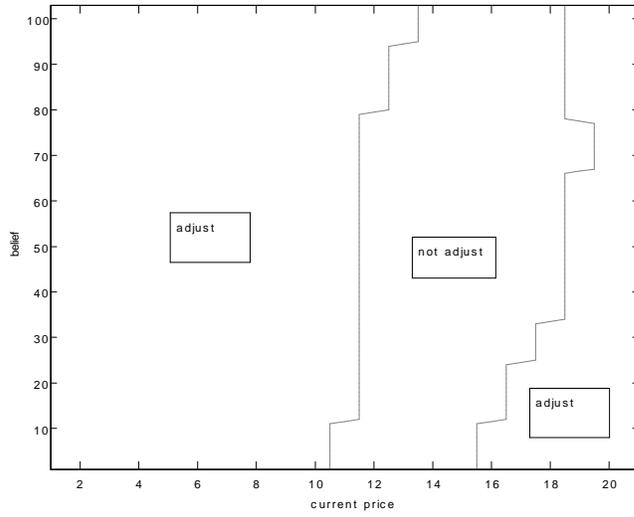


Figure 5: Optimal pricing policy.

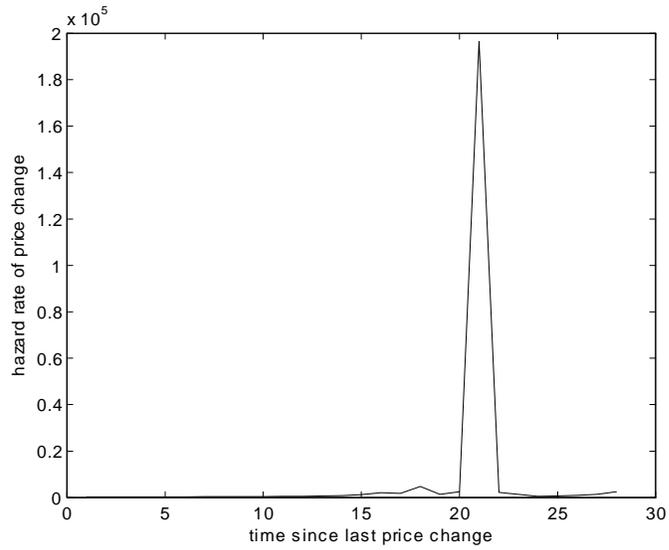


Figure 6: Hazard rate of price adjustment with menu costs

This exercise highlights a tension between data and models. In the data, nominal prices stay exactly constant for months at a time, strongly suggesting some form of menu costs. The same menu costs, however, imply that the prices that do change tend to be “old”, because far from the target. This is the essence of state-dependent pricing. In the data, however, most prices that change are young, they changed recently. It seems arduous to generate a declining hazard, as observed in the data, in a rational model without some form of learning. It must be the case that the act of changing a price alters the information set of the firm (which is what learning means), and that triggers a further change. If the information set did not change, the firm would not want to pay the menu cost twice in a row. Without menu costs, it would change prices all the time. We are currently exploring whether for extremely small menu costs the learning effect dominates at short durations and the menu cost at long durations, where prices are set under very accurate information.

9 Empirical evidence from CPI microdata

In the model without menu costs, price volatility should fall with time since entry (age of the firm), due to a standard selection effect common to all learning models. Firms that survive a long time must have been persuaded by a long series of strong sales to be facing a high level and a low elasticity of demand. Any additional piece of news has a negligible impact on this belief, either way, because it adds little to the previously accumulated evidence. As the optimal policy function maps directly beliefs into price, and is approximately linear at high beliefs, it follows that price volatility declines with firm age.

The shape of the optimal policy also suggests that all firms, even if old, should experience a burst in price volatility when they are in trouble, including when the economy is in a recession. The reason is twofold: firms experiment more with their price when close to giving up and leaving the market, thus generate more movement in beliefs, and the policy function has more curvature at low beliefs. Therefore, a sudden burst of price volatility should predict exit.

We explore empirically this correlation on monthly price observations on the Entry-Level Items (ELIs), such as “beverages”, that the Bureau of Labor Statistics collects to form the Consumer Price Index. We restrict attention to non-durable goods and services. Each ELI price is the average of many price observations in various stores for individual varieties of the ELI, such as “Coca Cola six-pack”. We identify a model firm with a single item product in the data. We specify a logit for the chance that the items exits, namely it is no longer found to be available for sale, as a function of (subsets of) the following variables: the standard deviation of price growth rates for that item over the previous four months (\hat{v}_4), which should predict exit;

the standard deviation of price growth rates for the ELI to which the item belongs since the ELI first appeared (\hat{v}), to control for the average variability of the price of that type of product or service; a monthly recession dummy (R), according to NBER’s business cycle dating, and its interaction with \hat{v}_4 ; the frequency of adjustment (f_4) of that item’s price in the last four months; and two controls relating to the 2-digit sector to which the ELI belongs, namely the sector-specific average CPI inflation rate (π_{SECT}) and average frequency of price adjustment (f_{SECT}). Finer sectorial classifications exist in the data but are difficult to implement for this specification. The data cover 1988-2010. Overall we have 2,971,175 item-month observations. For details, see Berger and Vavra (2010).

Table 1 reports the results. Standard errors are in parentheses. Estimates are very accurate thanks to the huge sample size. All specifications include controls at the sector level. In isolation, recent price volatility (\hat{v}_4) predicts a lower probability of exit (column I), the opposite of what the model indicates. The negative estimated coefficients on \hat{v}_4 in specifications I-II is due entirely (and more) to unobserved heterogeneity across items. When we control for the frequency of price changes of that item in the last four months (f_4) in (II), the estimated coefficients on \hat{v}_4 remains negative but is cut in half. When we control for the the overall volatility \hat{v} of price changes for the ELI to which the item belongs to since the beginning of the sample (or since the ELI entered the CPI basket, whichever happens later), the estimated coefficient on \hat{v} is so large and negative that the estimated coefficient of \hat{v}_4 turns positive, without (III) or with (IV) the cyclical interaction variable $\hat{v}_4 \times R$. In other words, items whose price is less sticky, in frequency or actual size of price changes, even controlling for the same characteristics at the sectorial level, are items that are much more likely to remain in the CPI for a longer time. When we control for unobserved heterogeneity with both f_4 and \hat{v} (V-VI), the positive effect of recent price volatility on exit grows even larger, and the cyclical interaction $\hat{v}_4 \times R$ preserves its positive sign. When we introduce the recession dummy R alone (VII-VIII), it has the expected strong positive effect on the probability of exit, and the interaction term with recent price volatility is no longer statistically significant, so most of the effect of the business cycle shows up as a direct impact. But the strong predictive power of recent idiosyncratic price volatility on exit is robust and, if anything, stronger.

The reason why the two measures of price stickiness, frequency f_4 and size \hat{v} of log price changes, play different roles presumably is that the latter may be large also with rare, but huge, price changes. All estimated coefficients, except as discussed that of \hat{v}_4 , remain quite stable in size across specifications. To interpret the size of an estimated logit coefficient, for example, in column VI, when the standard deviation of log price changes \hat{v}_4 goes up by one percentage point, the relative likelihood of forced substitution of the item goes up by multiplicative factor

of $e^{2.85} - 1 = 16.288$.

The empirical literature on nominal price rigidity that exploits the same CPI data has widely documented vast cross-item heterogeneity in price stickiness. Monthly price volatility ranges from very high for items like gasoline, whose price change almost daily, to very low for prices that change every year or two, like magazines'. The results suggest the following interpretation. Items that have unusually high price volatility \hat{v}_4 in recent months (relative to their average experience \hat{v}) are more likely to exit. This is a robust feature of the data, which is very naturally explained by our learning-and-selection model, and is at least in part due to an aging effect. It is well known that older firms are more productive and less volatile. Similarly, older CPI items must be particularly valuable to consumers, because they survived spending decisions for a long time. Therefore, consumers learned that these items are desirable, their prices are stable as new sales do not reveal much new about the market tastes for that item. As many ELIs are added to the CPI basket long after first appearing in the market, the data contain a left-censored age of the item, which does not allow us to directly control for it. We conjecture that movements in \hat{v}_4 given \hat{v} capture in part age. Overall, we see this empirical evidence—that unusually high recent price volatility leads exit of the item from the market—as consistent with our theory.

	I	II	III	IV	V	VI	VII	VIII
\hat{v}_4	-2.46	-1.00	1.78	1.37	3.18	2.85	3.17	3.22
	(.09)	(.10)	(.11)	(.12)	(.11)	(.12)	(.11)	(.11)
\hat{v}			-20.44	-20.42	-20.12	-20.11	-20.05	-20.05
			(.34)	(.34)	(.34)	(.34)	(.34)	(.34)
R							.41	.42
							(.014)	(.015)
$\hat{v}_4 \times R$				1.97		1.52		-.31
				(.19)		(.16)		(.21)
f_4		-1.05			-1.12	-1.10	-1.12	-1.12
		(.05)			(.05)	(.05)	(.05)	(.05)
π_{SECT}	-.10	-.11	-.09	-.09	-.10	-.10	-.10	-.10
	(.006)	(.006)	(.006)	(.006)	(.006)	(.006)	(.006)	(.006)
f_{SECT}	-3.32	-2.93	-1.97	-1.98	-1.53	-1.55	-1.55	-1.55
	(.06)	(.06)	(.05)	(.06)	(.06)	(.06)	(.06)	(.06)
pseudo R^2	.023	.0235	.0371	.0373	.0385	.0395	.04	.04

Table 1. Logit estimates for the chance of forced substitution of a CPI item

10 Extensions and conclusions

It is plausible that the demand elasticity σ changes stochastically, unobserved by the firm (other than through noisy sales). We envision the following process. New varieties are introduced each period. For example, the iPhone, when first introduced, may have been considered similar to a Blackberry. Sales (and stock market) then revealed that this was not the case, and the iPhone was quite a unique device. New smartphones introduced later, however, may or may not be close substitutes for the iPhone. Only sales data will tell. Therefore the true elasticity σ has a positive drift, which may depend on the amount of entry. Given the Kalman-filter-like structure of our observation equation (5), this extension is straightforward to implement. It describes well a product life cycle.

A Appendix

A.1 Proof of Proposition 1

The myopic optimal price maximizes expected static profits, that we denote by Π^{MY} . We show that Π^{MY} is supermodular in P, λ at least near the optimizer P^{MY} . Taking a derivative w.r. to λ

$$\begin{aligned} \frac{\partial \Pi^{MY}(P, \lambda)}{\partial \lambda} &= \left(P - \frac{w}{z}\right) \frac{Y}{2} \left[\frac{P^{-\sigma_1}}{\bar{P}_1^{1-\sigma_1}} - \frac{P^{-\sigma_2}}{\bar{P}_2^{1-\sigma_2}} \right] \\ &= \frac{\Pi^{MY}(P, \lambda) + \Psi - \left(P - \frac{w}{z}\right) \frac{P^{-\sigma_2}}{\bar{P}_2^{1-\sigma_2}} Y/2}{\lambda Y/2} = \frac{2}{\lambda Y} [\Pi^{MY}(P, \lambda) - \Pi^{MY}(P, 0)] \end{aligned}$$

Supermodularity of $\Pi^{MY}(P, \lambda)$ is equivalent to this function being increasing in P at $P = P^{MY}(\lambda)$. Taking another derivative w.r. to P , evaluating it at $P^{MY}(\lambda)$ and using the fact that, by definition, $P^{MY}(\lambda)$ maximizes $\Pi^{MY}(P, \lambda)$,

$$\left[\frac{\partial^2 \Pi^{MY}(P, \lambda)}{\partial \lambda \partial P} \right]_{P=P^{MY}(\lambda)} = -\frac{2}{\lambda Y} \frac{d}{dP} [\Pi^{MY}(P, 0)]_{P=P^{MY}(\lambda)}$$

By concavity of $\Pi^{MY}(P, 0)$ in P , this is positive if and only if the optimal price exceeds the one chosen under perfect information at $\lambda = 0$: $P^{MY}(\lambda) > P^{MY}(0) = P_2^{PI}$. We draw the following conclusion:

$$\frac{dP^{MY}(\lambda)}{d\lambda} > 0 \Leftrightarrow P^{MY}(\lambda) > P^{MY}(0).$$

The last inequality is true at $\lambda = 1$ because $P^{MY}(1) = P_1^{PI} > P_2^{PI} = P^{MY}(0)$ by $\sigma_2 > \sigma_1$. By continuity, it is true for λ large enough. By contradiction, if for some $\lambda' > 0$ we have $P^{MY}(\lambda') = P^{MY}(0)$, then either $P^{MY}(\lambda' - \Delta) > P^{MY}(0)$ for $\Delta > 0$ small enough, but

this requires $\frac{dP^{MY}(\lambda' - \Delta)}{d\lambda} < 0$, which contradicts the equivalence above; or $P^{MY}(\lambda' - \Delta) = P^{MY}(0)$ over an open interval of small positive values of Δ , but this equality cannot hold on an interval by the strict concavity of expected profits; or $P^{MY}(\lambda' - \Delta) < P^{MY}(0)$, but then $\frac{dP^{MY}(\lambda')}{d\lambda} < 0$ for all $\lambda \in (0, \lambda')$, thus $P^{MY}(\lambda)$ keeps declining with $\lambda \downarrow 0$ and stays below $P^{MY}(0)$, until $P^{MY}(0+) < P^{MY}(0)$. But, by the Theorem of the Maximum, the optimizer $P^{MY}(\cdot)$ is continuous, which yields the final, desired contradiction.

A.2 Bayesian updating with log-Normal preference shocks

Assume $\varepsilon = \log \alpha \sim N(m, s^2)$. The log-Likelihood Ratio (log-LR) of state 2 to state 1 is, in state $i = 1$ (when $q = \mu_1 - \sigma_1 p + y + \varepsilon$),

$$\frac{(\varepsilon - m)^2 - (\mu_1 - \sigma_1 p + \varepsilon - \mu_2 + \sigma_2 p - m)^2}{2s^2} = \frac{\hat{\zeta}^2 - \left(\hat{\zeta} - \Delta\mu + p\Delta\sigma\right)^2}{2s^2}$$

where $\hat{\zeta} = \varepsilon - m \sim N(0, s^2)$, and

$$\begin{aligned}\Delta\mu &= \mu_2 - \mu_1 \\ \Delta\sigma &= \sigma_2 - \sigma_1\end{aligned}$$

Similarly, in state 2 (when $q = \mu_2 - \sigma_2 p + y + \varepsilon$) the log-LR is

$$\begin{aligned}& \frac{(\mu_2 - \sigma_2 p + \varepsilon - \mu_1 + \sigma_1 p - m)^2 - (\varepsilon - m)^2}{2s^2} \\ &= \frac{\left(\Delta\mu - p\Delta\sigma + \hat{\zeta}\right)^2 - \hat{\zeta}^2}{2s^2} = \frac{\left(-\hat{\zeta} - \Delta\mu + p\Delta\sigma\right)^2 - \hat{\zeta}^2}{2s^2}\end{aligned}$$

Let

$$\tilde{\zeta} = -\hat{\zeta} = m_\varepsilon - \varepsilon \sim N(0, s^2)$$

so the log-LR in state 2 is

$$\frac{\left(\tilde{\zeta} - \Delta\mu + p\Delta\sigma\right)^2 - \left(-\tilde{\zeta}\right)^2}{2s^2} = \frac{\left(\tilde{\zeta} - \Delta\mu + p\Delta\sigma\right)^2 - \tilde{\zeta}^2}{2s^2}.$$

If we let $\zeta \sim N(0, s^2)$ denote a generic random variable independent of the state, the log-LR in each state is the same random variable except that the state i causes a sign switch:

$$(-1)^{\mathbb{I}(i=2)} \frac{\zeta^2 - (\Delta\sigma p - \Delta\mu + \zeta)^2}{2s^2}.$$

where \mathbb{I} is the indicator function. That is, $\tilde{\zeta}$ and $\hat{\zeta}$ have the same distribution, independent of the state, that we capture through ζ .

References

- Alvarez, F., F. Lippi and L. Paciello, 2011, “Optimal price setting with observation and menu costs”, *Quarterly Journal of Economics*, 126, pp. 1909-1960.
- Arellano, C., Y. Bai, P. Kehoe, 2012, “Financial Markets and Fluctuations in Uncertainty”, Minneapolis FED Research Department Staff Report 466.
- Bachmann, R. and C. Bayer, 2011a, “Investment Dispersion and the Business Cycle”, NBER WP 16861.
- Bachmann, R. and C. Bayer, 2011b, “Uncertainty Business Cycles - Really?”, NBER WP 16862.
- Bachmann, R., S. Elstner, and E. Sims, 2010, “Uncertainty and Economic Activity: Evidence from Business Survey Data”, NBER WP 16143. Forthcoming in the *American Economic Journal: Macroeconomics*.
- Baker, Scott, Nicholas Bloom, and Steven J. Davis. 2012. “Measuring Economic Policy Uncertainty”, mimeo, Stanford University.
- Berger, D. and J. Vavra, 2010, “Dynamics of the U.S. Price Distribution”, mimeo NWU and U Chicago.
- Bloom, N., 2009, “The Impact of Uncertainty Shocks”, *Econometrica*, May.
- Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, S. Terry, 2012, “Really Uncertain Business Cycles”, NBER WP 18245.
- Born, B. and J. Pfeifer, 2012, “Policy Risk and the Business Cycle”, mimeo U Mannheim.
- Brainard, L. and D. Cutler, 1993, “Sectoral Shifts and Cyclical Unemployment Revisited”, *Quarterly Journal of Economics*.
- Campbell J.Y, M. Lettau, B. Malkiel and Y. Xu, 2001, “Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk”, *Journal of Finance* 56(1), 1-41.
- Cui, W. 2012, “Delayed Capital Reallocation”, mimeo Princeton University.
- Campbell, J. R. and B. Eden, 2010, “Rigid Prices: Evidence from U.S. Scanner Data”, July 2010, WP 2005-08
- Davis, S., J. Haltiwanger and S. Schuh, 1996, *Job Creation and Destruction*, MIT Press.
- D’Erasmus, P. and H. Moscoso Boedo, 2012, “Intangibles and Endogenous Firm Volatility over the Business Cycle”. mimeo University of Virginia
- Di Tella, S., 2012, “Uncertainty Shocks and Balance Sheet Recessions”, mimeo MIT.
- Fernandez-Villaverde, J., P. Guerron-Quintana, J. Rubio-Ramirez, and Martin Uribe, 2011, “Risk Matters: The Real Effects of Stochastic Volatility Shocks”, *American Economic Review*, 101: 2530-61.

Fernandez-Villaverde, J., P. Guerron-Quintana, K. Kuester, J. Rubio-Ramirez, 2012 “Fiscal Volatility Shocks and Economic Activity”, mimeo.

Gilchrist, S., V. Yankov and E. Zakrajsek, 2009, “Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets”, *Journal of Monetary Economics*, 56(4), pp. 471-493.

Johannsen, B., 2012, “When are the Effects of Fiscal Policy Uncertainty Large?”, mimeo NWU.

Kotler, P., 2000, *Marketing Management*, The Millenium Edition, Prentice Hall, Upper Saddle River, NJ.

Lilien, D.M., 1982, “Sectoral shifts and cyclical unemployment”, *Journal of Political Economy*, 90(4), 777-793

Nakamura, E. and J. Steinsson, 2008, “Five Facts About Prices: A Reevaluation of Menu Cost Models”, *Quarterly Journal of Economics*, 123(4), 1415-1464.

Trefler, D., 1993, “The Ignorant Monopolist: Optimal Learning with Endogenous Information,” *International Economic Review*.

Vavra, J., 2010, “The Empirical Price Duration Distribution and Monetary Non-Neutrality”, mimeo.

Vavra, J., 2012, “Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation”, mimeo.