FIRM UNCERTAINTY CYCLES AND THE PROPAGATION OF NOMINAL SHOCKS*

Isaac Baley† **Julio A. Blanco** ‡

October 18, 2016

Abstract

Firms operate in constantly changing and uncertain environments. We argue that firm uncertainty is a key determinant of pricing decisions, and that it affects the propagation of nominal shocks in the economy. For this purpose, we develop a price-setting model with menu costs and imperfect information about idiosyncratic productivity. Uncertainty arises from firms' inability to distinguish between permanent and transitory productivity changes. Upon the arrival of a productivity shock, a firm's uncertainty spikes up and then fades in light of new information until the next shock arrives. These idiosyncratic uncertainty cycles, when paired with menu costs, generate endogenous price flexibility that correlates positively with uncertainty. When heterogeneity in firm uncertainty is disciplined with micro-price statistics, aggregate nominal shocks have very persistent effects on output. However, if nominal shocks are accompanied by an increase in the average level of uncertainty, their output effects are reduced.

JEL: D80, E30, E50

Keywords: Menu costs, firm uncertainty, information frictions, monetary policy.

*Previously circulated as "Menu Costs, Uncertainty Cycles, and the Propagation of Nominal Shocks." We are especially thankful to Virgiliu Midrigan and Laura Veldkamp for their advice and to anonymous referees for their constructive comments. We also thank Fernando Álvarez, Rudi Bachmann, Anmol Bhandari, Jarda Borovička, Katka Borovičková, Olivier Coibion, Mark Gertler, Ricardo Lagos, John Leahy, Francesco Lippi, Robert E. Lucas, Rody Manuelli, Cynthia-Marie Marmo, Simon Mongey, Joseph Mullins, Emi Nakamura, Gastón Navarro, Ricardo Reis, Tom Sargent, Edouard Schaal, Ennio Stacchetti, Venky Venkateswaran, Jaume Ventura, as well as seminar participants at 4th Ifo Conference on Macroeconomics and Survey Data 2013, Midwest Economics Association 2013, SED Meetings 2013, ASSA Meetings 2015, Stanford Institute for Theoretical Economics 2015, Econometric Society Meetings 2015, 40 Simposio de la Asociación Española de Economía, Barcelona GSE Summer Forum 2016, New York University, NYU Stern, Princeton, Washington University St. Louis, St. Louis Fed, Federal Reserve Board, University of Toronto, Einaudi Institute, CREI, Universitat Pompeu Fabra, BIS, Singapore Management University, Carnegie Mellon, UC Davis, University of Melbourne, University of Sydney, Banco de México, ITAM, Oxford, and Universitat Autònoma de Barcelona for very useful comments and suggestions. Julio A. Blanco gratefully acknowledges the hospitality of the St. Louis Fed where part of this paper was completed.

†Universitat Pompeu Fabra and Barcelona GSE, isaac.baley@upf.edu

‡University of Michigan, jablanco@umich.edu

1 Introduction

Firms operate in constantly changing environments. Fresh technologies become available, new products are developed, unfamiliar markets and competitors appear, workers are replaced, and supply chains get disrupted. These idiosyncratic changes are recurrent, large, and often permanent; and many times, firms do not have all the information needed to assess their effects. The lack of perfect knowledge generates uncertainty that affects firms' actions, and in particular, their pricing decisions. In this context, many questions arise: How do firms learn about their environment and respond to its changes? Does uncertainty increase or decrease price flexibility? Does uncertainty heterogeneity matter for the propagation of nominal shocks? Is it relevant for monetary policy?

In this paper we argue that firm idiosyncratic uncertainty is a key determinant of pricing decisions, and as a consequence, it shapes how nominal shocks propagate and affect output in the economy. We show that more uncertain firms have more flexible pricing rules and respond faster to changes in their environment compared with more certain firms. As a result, the flexibility of the aggregate price level depends on the dispersion of uncertainty across firms. When heterogeneity in firm uncertainty is disciplined with micro-price statistics, the aggregate price level is more rigid than in an economy with one type of firm; thus nominal shocks have larger and more persistent effects on output. However, if nominal shocks are accompanied by an increase in the average level of uncertainty, their output effects are smaller as firms prices become more responsive. Our results highlight the importance of taking into account firm uncertainty, and specially its cross-sectional distribution, to assess the effect of a monetary shock; in this way, firm uncertainty becomes relevant for policy making decisions.

To obtain these results, we build a price-setting model that involves nominal and informational frictions. Firms face a menu cost to adjust their prices and are uncertain about their level of productivity. In particular, we assume that the firms receive permanent and transitory shocks to their idiosyncratic productivity, but they cannot distinguish between types of shocks. Because firms must pay a menu cost with each adjustment, it is optimal to ignore transitory shocks and only respond to permanent shocks. Firms follow a Jovanovic (1979) type of gradual learning using Bayes' law estimate the permanent component of their productivity. We call the conditional variance of the estimates firm uncertainty. As in any problem with fixed adjustment costs, the decision rule takes the form of an inaction region, in which the firm adjusts her price only if she receives shocks that make it worth paying the menu cost. In this case, the inaction region also depends on firm uncertainty. A new insight from this paper is the fact that inaction regions refer to estimates about the individual state, and not the true state. This makes a difference because after a firm takes action, her judgement might turn out to be wrong, thus leading her to take action again very soon.

One of our framework's innovations is a structure of productivity shocks that gives rise to idiosyncratic uncertainty cycles, defined as recurrent episodes of high uncertainty followed by episodes of low uncertainty at the firm level. The key to generate these cycles are infrequent and large shocks to permanent idiosyncratic productivity—or regime changes—where the timing but not the magnitude of the shock is perfectly known.[1] That is, a firm knows when a regime change has occurred, but she does not have perfect information about the magnitude of the change. When a regime change shock hits, uncertainty spikes up; then it fades with learning until it jumps again with the arrival of the next shock; these are the uncertainty cycles. The interaction of uncertainty cycles with inaction regions that depend on uncertainty generates heterogeneity in price flexibility in ex-ante identical firms. This heterogeneity in price flexibility plays two key roles. On the one hand, heterogeneity generates a decreasing hazard rate of price adjustment that allows us back out the distribution of firm uncertainty from micro-price data. On the other hand, heterogeneity amplifies nominal rigidities by delaying the price response of low uncertainty firms.

[1] Large and infrequent idiosyncratic shocks to productivity were first introduced in menu cost models by Gertler and Leahy (2008), and then used by Midrigan (2011) as a way to account for the empirical patterns of pricing behavior, such as fat tails in price change distributions. In our model, the infrequent first moment shocks paired with the information friction give rise to second moment shocks in beliefs, or uncertainty shocks.

The regime change shocks are crucial to produce a non-degenerate distribution of uncertainty that keeps heterogeneity active in steady state. Without regime changes, uncertainty becomes constant and equal across firms in steady state. Álvarez, Lippi and Paciello (2011) study a related problem where firms pay menu costs for price adjustment and observation costs to learn about their continuous state. In the particular case of infinite observation costs, firms receive noisy signals about their state. In the absence of regime changes, uncertainty stabilizes around a constant value and all firms behave as in the standard menu cost model of Golosov and Lucas (2007) where there is large monetary neutrality. The novelty here is that such stabilization is prevented by the regime changes, heterogeneity persists in steady state, and monetary neutrality is diminished.

We continue with an overview of the ideas in this paper and how they relate with other contributions in the literature.

Uncertainty, inaction regions, and decreasing hazard Our theoretical contribution is twofold. First, we contribute to the filtering literature by extending the Kalman-Bucy filter to an environment where the state follows a general jump-diffusion process. Second, we characterize analytically the dynamic inaction region and several price statistics as a function of uncertainty. This involves solving a stopping time problem together with a signal extraction problem. This analytical characterization allows for understanding how uncertainty shapes pricing decisions. The model is very general and it is easily extendable to a variety of environments that involve non-convex adjustment costs and idiosyncratic uncertainty shocks. For example, Senga (2016) uses of a similar mechanism in a model of investment and misallocation, in which firms occasionally experience a shock that forces them to start learning afresh about their productivity.

The mechanism that generates a decreasing hazard rate comes from the combination of the uncertainty cycles and a positive relationship between uncertainty and adjustment frequency. This positive relationship is subtle, as uncertainty has two opposing effects on frequency. Higher uncertainty means that the firm does not trust her current estimates of permanent productivity, and thus she optimally puts a high Bayesian weight on her observations, which are random by construction. Estimates become more volatile and the probability of leaving the inaction region and adjusting the price increases. This is known as the "volatility effect" and it has a positive effect on the adjustment frequency. This volatility arises from belief uncertainty. As a reaction to the volatility effect, which triggers more price changes and menu costs payments, the optimal policy calls for saving menu costs by widening the inaction region. This is known as "option value effect" (Barro (1972) and Dixit (1991)), and it has a negative effect on the adjustment frequency. However, the widening of the inaction region does not compensate for the increase in volatility. Overall, the volatility effect dominates and higher uncertainty yields higher adjustment frequency. When this relationship is paired with uncertainty cycles, we obtain adjustment frequency cycles as well: firms alternate between periods of high frequency with periods of low frequency; in other words, price changes get clustered in some periods instead of evenly spread across time. This gives rise to the decreasing hazard rate of price adjustment.

With respect to the positive relationship between uncertainty and adjustment frequency, Bachmann, Born, Elstner and Grimme (2013) use survey data collected from German firms to document a positive relationship between the variance of firm-specific forecast errors on sales—a measure of firm-level belief uncertainty—and the individual adjustment frequency. Vavra (2014) and Karadi and Reiff (2014) exploit a version of this positive relationship in menu cost models where productivity shocks volatility follows exogenous autoregressive processes. Both belief uncertainty and fundamental volatility shocks generate higher adjustment frequency in a menu cost model. However, decreasing hazards cannot be generated by autoregressive processes, the jumps are needed.

Regarding decreasing hazard rates of price adjustment, other alternative explanations are discounts in Kehoe and Midrigan (2015), mean reverting shocks in Nakamura and Steinsson (2008), experimentation in Bachmann and Moscarini (2011), introduction of new products in Argente and Yeh (2015), price plans in Álvarez and Lippi (2015), and rational inattention in Matějka (2015). Empirically, decreasing hazards are documented in several datasets, covering different countries and different periods. For instance, decreasing hazards are documented by Nakamura and Steinsson (2008) using monthly BLS data for consumer and producer prices, Eden and Jaremski (2009) using Dominick's weekly scanner data, Dhyne *et al.* (2006) using monthly CPI data for Euro zone countries, and Cortés, Murillo and Ramos-Francia (2012) for CPI data in Mexico. These papers control for observed and unobserved heterogeneity by imposing structure on the heterogeneity across items and also filter discounts out; these are known sources of potential downward bias in the hazard rates' slope. Vavra (2010) and Campbell and Eden (2014) also find downward sloping structural duration dependence by estimating within-item hazards rather than pooling across items; the first uses CPI and Dominick's data, and the second uses retailer scanner data. We propose a new methodology that controls for heterogeneity in adjustment frequency and eliminates survivor bias when estimating hazard rates. Our method uses the relative stopping times distribution, which are the stopping times normalized by the average duration of an item's price. Using disaggregated item-level CPI data from the UK, to which we apply discount filters and other standard procedures, we also document a decreasing hazard. In spite of all the previous results and our own empirical finding, Klenow and Kryvtsov (2008) find a flat hazard for CPI data when controlling for frequency deciles. Since the evidence is not conclusive, we provide additional support for the our theory using cross-sectional implications of our learning model, such as age-dependent price statistics.

Age dependent pricing An interesting prediction of our learning model is that price age, defined as the time elapsed since its last change, is a determinant of the size and frequency of its next adjustment. Young prices—or recently set, mostly by firms who are highly uncertain at the time of the change—and old prices—set many periods ago by firms which are currently certain about their productivity– exhibit different behavior. In particular, young prices are more likely to be reset than older prices. Furthermore, as the inaction region decreases with uncertainty and price age, young prices changes will tend to be larger and more dispersed compared to older prices. These predictions are documented by Campbell and Eden (2014) using weekly scanner data for the retail sector. It finds that young prices (set less than three weeks ago) are relatively more dispersed and more likely to be reset than older prices. Further evidence regarding age dependence is documented in Baley, Kochen and Sámano (2016), which uses comprehensive item-level Mexican CPI data at weekly frequency to document that adjustment frequency and price change dispersion fall with the age of the price.

Decreasing hazard and propagation of monetary shocks Why does a decreasing hazard rate imply more persistent monetary shock effects on output? To answer this question, it is key to recognize two observations. First, a decreasing hazard rate generates cross-sectional heterogeneity. At the firm level, a falling hazard is equivalent to having time-varying adjustment frequency; in the aggregate, it implies that there are different types of firms: high frequency firms and low frequency firms. Second, a firm's *first* price change after a monetary shock takes care of incorporating the monetary shock into her price and, in the absence of complementarities, it is the only price change that matters for the accounting of monetary effects. Any price changes after the first one are the result of idiosyncratic shocks that cancel out in the aggregate and do not contribute to changes in the aggregate price level. When a monetary shock arrives, the high frequency firms will incorporate almost immediately the monetary shock with their first price change; but the monetary shock will have effects until the low frequency firms have made their first price adjustment. Therefore, the heterogeneity generated by a decreasing hazard makes the aggregate price level less responsive to monetary shocks compared to an aggregate price level where every firm faces the same average frequency. Amplification of monetary non-neutrality due to dispersion of times until the first adjustment is discussed in Carvalho and Schwartzman (2015) and Álvarez, Lippi and Paciello (2016) for time-dependent models.

Heterogeneity in adjustment frequency has been analyzed as a source of non-neutrality before. For instance, Carvalho (2006) and Nakamura and Steinsson (2010) find larger non-neutralities in sticky price models with exogenous heterogeneity in sector level adjustment frequency. Heterogeneity in our setup arises endogenously in ex-ante identical firms that churn between high and low levels of uncertainty. Importantly, this type of heterogeneity does not refer to different types of firms, but to different uncertainty states within each firm. Therefore, our mechanism does not rely on survivor bias to generate a decreasing hazard.[2]

The following simplified example highlights the main mechanisms in our framework. Suppose there is a continuum of firms and two states for uncertainty, high and low; assume that half of the firms are in each state. High uncertainty firms change their price during N consecutive periods and then become low uncertainty firms with probability one; this switch in firm type captures the learning process. Low uncertainty firms do not change their price and with probability $1/N$ they become high uncertainty firms; this switch in firm type captures the regime changes. In steady state, the aggregate adjustment frequency is equal to $1/2$. Now suppose there is a monetary shock. To measure the output effects, let us keep track of the mass of firms that have not adjusted their price. On impact, $1/2$ of the firms (all high uncertainty firms) change their price and the output effect is equal to $1/2$ (all low uncertainty firms). In subsequent periods, all high uncertainty firms adjust again, but we do not count these price changes towards the effect of the monetary shock because these respond only to idiosyncratic shocks. Then the low uncertainty firms that become high uncertainty (a fraction $1/N$ of firms) adjust and incorporate the monetary shock. Therefore, the output effect is $1/2(1-1/N)$, which is equal to the mass of low uncertainty firms that have not switched yet. Continuing in this way, the output effect τ periods after the impact of the monetary shock is given by $1/2(1-1/N)^\tau$. The persistence of the output response is driven by N, which is the number of periods that firms remain characterized by high uncertainty (the speed of learning). Now let us compare this stylized economy with learning to a Calvo economy with the same aggregate frequency, which is generated with a random probability of adjustment of $1/2$. On impact, the output effects also equal to $1/2$, but in subsequent periods the response is $1/2(1-1/2)^\tau$. Therefore, as long as $N > 2$, the economy with learning has more persistence than the Calvo economy.

Larger persistence of output response to monetary shocks To give a quantitative assessment of the impact of monetary shocks implied by the model, we study a general equilibrium economy with a continuum of firms that solve the price-setting problem with menu costs and idiosyncratic uncertainty cycles. It is a Bewley–type model with ex-ante identical firms who are different ex-post. The environment also includes a representative household that provides labor in exchange for a wage, consumes a bundle of goods produced by the firms, and holds real money balances. We solve for the steady state of this economy and calibrate the parameters to match UK micro price statistics computed by us. We target three factors jointly: the average adjustment frequency, the dispersion of the price change distribution, and the decreasing hazard rate. In particular, we use the hazard rate slope to calibrate the volatility of the transitory shocks that give rise to the information friction. This approach of using a price statistic to recover information parameters was first suggested in Jovanovic (1979), and Borovičková (2013) uses it to calibrate a signal-noise ratio in a labor market framework.

In the calibrated economy we study the effect of a small unanticipated increase in the money supply. In equilibrium this monetary shock increases wages and gives incentives to firms to increase their prices. As a baseline case, we assume that the monetary shock is perfectly observable and then relax this assumption. The results show that the output response to the monetary shock is more persistent in our model than in alternative models. The larger persistence generated in the baseline model only relies on information frictions regarding idiosyncratic conditions; the arrival of the aggregate nominal shock is perfectly observed by firms.

The model performs well in terms of the long-run effects of the monetary shock by increasing persistence, but it has shortcomings with respect to its short-run response. On impact of the monetary shock, the adjustment frequency overshoots as a result of a large mass of firms with low uncertainty and small inaction regions. However,

[2] Survivor bias emerges when computing hazards in populations with heterogenous types as noted by Kiefer (1988) and studied in an economy with different Calvo agents as in Álvarez, Burriel and Hernando (2005).

this overshoot is not observed in the data. Blanco (2016b) also finds this overshoot in a full-fledged menu cost DSGE model with zero lower bound that is coherent with micro-price statistics and business cycle facts.

To address this issue, we consider an extension of the model that incorporates an additional information friction. We assume that the monetary shock is only partially observed by firms. This type of constraint on the information set regarding aggregate shocks are at the core of the pricing literature with information frictions that started with Lucas (1972) and has been recently explored by Mankiw and Reis (2002), Woodford (2009), Maćkowiak and Wiederholt (2009), Hellwig and Venkateswaran (2009), and Álvarez, Lippi and Paciello (2011), among others. These firms apply the same learning technology to filter the monetary shock as they do to filter their idiosyncratic permanent productivity shocks. Upon the impact of the monetary shock, there will be initial forecast errors that disappear over time. The persistence of forecast errors increases the persistence of the output response. Under this assumption, the output response is significantly amplified compared to the case with the observable monetary shock.

Aggregate uncertainty, forecast errors, and persistence The model also predicts that unobserved monetary shocks have smaller effects when aggregate uncertainty is high. We interact the monetary shock with a synchronized uncertainty shock across all firms. In more uncertain times, firms place a higher weight on new information, forecast errors disappear faster, and the monetary shock is quickly incorporated into prices; this reduces the persistence of the average forecast error, and in turn, the persistence of the output response. This relationship between uncertainty and forecast errors is novel and there is empirical evidence in this respect. For instance, Coibion and Gorodnichenko (2015) compares the dynamics of forecast errors during periods of high economic volatility (such as the 70's and 80's) with periods of low economic volatility (such as the late 90's). It concludes that information rigidities are higher during periods of low uncertainty than higher uncertainty. The joint dynamics of uncertainty, prices, and forecast errors implied by our model provide a theoretical framework to think about this piece of evidence. Furthermore, we show how forecast errors can be disciplined with micro-price data.

The negative relationship between the effects of monetary shocks and aggregate uncertainty is also documented empirically in various studies. Pellegrino (2014) finds weaker real effects of monetary policy shocks during periods of high uncertainty, and even more, it finds that prices respond more to a monetary shock during times of greater firm-level uncertainty. Aastveit, Natvik and Sola (2013) shows that monetary shocks produce less output effects when various measures of economic uncertainty are high; and other papers find differential effects of monetary shocks in good and bad times, where bad times are associated with periods of high uncertainty, as Caggiano, Castelnuovo and Nodari (2014), Tenreyro and Thwaites (2015), Mumtaz and Surico (2015). Finally, Vavra (2014) uses BLS data to document that the cross-sectional dispersion of price changes (a measure of aggregate uncertainty) is larger during recessions, implying higher price level flexibility and lower effects of monetary policy.

Uncertainty and Passthrough The previous results concern the response of the aggregate price level to a monetary shock. To examine the responsiveness of individual prices, we follow the methodology used to estimate the exchange rate pass-though as in Gopinath, Itskhoki and Rigobón (2010). We consider stochastic money supply and simulate a panel of firms. Then we regress the size of price changes on the cumulative monetary shock between price changes to obtain the medium-run pass-through coefficient. We find that with observable monetary shocks, pass-through is complete: when prices adjust, they fully incorporate the money shock. When the money shock is unobserved, pass-through is five times smaller, as suggested by empirical studies. Our contribution to this literature lies in showing that a menu cost model with information frictions that is coherent with micro-price statistics can reduce nominal pass-through. Furthermore, we show that idiosyncratic uncertainty can generate a positive relationship between the standard deviation of price changes and pass-through, as documented in Berger and Vavra (2015) in the context of import price-setting.

2 Firm problem with nominal rigidities and information frictions

We develop a model that combines an inaction problem arising from a non-convex adjustment cost together with a signal extraction problem. Although the focus here is on pricing decisions, the model is easy to generalize to other settings. We contribute in two ways. First, we provide filtering equations for a state that has both continuous and jump processes. Second, we derive closed form decision rules that take the form of a time-varying inaction region that reflects the uncertainty dynamics.

2.1 Environment

Consider a profit maximizing firm that chooses the price at which to sell her product, subject to idiosyncratic productivity (or cost) shocks. She must pay a menu cost θ in units of product every time she changes the price. We assume that in the absence of the menu cost, the firm would like to set a price that makes her markup—price over marginal cost—constant. The productivity shocks—and therefore her markup—are not perfectly observed, only noisy signals are available to the firm[3]. She chooses the timing of the adjustments as well as the new reset markups. Time is continuous and the firm discounts the future at a rate r.

Quadratic loss function Let μ_t be the markup gap, defined as the log difference between the current markup and the optimal markup obtained from a static problem without menu costs. Firms incur an instantaneous quadratic loss as the markup gap moves away from zero:

$$\Pi(\mu_t) = -B\mu_t^2, \qquad B > 0$$

Quadratic profit functions are standard in price setting models, such as Barro (1972) and Caplin and Leahy (1997), and can be motivated as second order approximations of more general profit functions.

Markup gap process The markup gap μ_t follows a jump-diffusion process as in Merton (1976)

$$d\mu_t = \sigma_f dW_t + \sigma_u u_t dQ_t \tag{1}$$

where W_t is a Wiener process, $u_t Q_t$ is a compound Poisson process with the Poisson counter's intensity λ, and σ_f and σ_u are the respective volatilities. When $dQ_t = 1$, the markup gap receives a Gaussian innovation $u_t \sim \mathcal{N}(0,1)$. The process Q_t is independent of W_t and u_t. This process for markup gaps nests two specifications that are benchmarks in the literature:

i) small *frequent* shocks modeled as the Wiener process W_t with small volatility σ_f; these shocks are the driving force in standard menu cost models, such as Golosov and Lucas (2007)[4];

ii) large *infrequent* shocks modeled through the Poisson process Q_t with large volatility σ_u. These shocks produce a leptokurtic distribution of price changes and are used in Gertler and Leahy (2008) and Midrigan (2011) to capture the fat tailed price change distribution in the data.

[3] In Álvarez, Lippi and Paciello (2011) firms pay an observation cost to see their true productivity level; here we make the observation cost infinite and the true state is never fully revealed. The Appendix of that paper discusses this particular case in an environment where the information friction does not have effects in steady state.

[4] Golosov and Lucas (2007) use a mean reverting process for productivity instead of a random walk. Still, our results concerning small frequent shocks will be compared with their setup.

Two remarks on the markup process We think of markup fluctuations as the result of idiosyncratic productivity or cost shocks, but the setup is flexible enough to allow for alternative interpretations. For instance, if firm's demand function comes from a Dixit-Stiglitz structure as in the general equilibrium model of Section 4, fluctuations in costs are isomorphic to fluctuations in the demand elasticity, as both shocks enter markup gaps in the same way. While the interpretation of imperfect information on the demand structure might be more adequate in some applications, the effects on markups are the same under either assumption and the results do not change.

When we calibrate the markup process to match micro price statistics, we find that the volatility of infrequent shocks σ_u is very large relative to the volatility of frequent shocks σ_f (see Section 4.3 for details). This parametrization breaks the Normality of markup growth and generates a leptokurtic —or fat tailed— price change distribution. Prices are not the only firm outcomes that display this behavior. Section A of the Online Appendix documents leptokurtic distributions for profit, employment, sales, and capital growth rates for firms in COMPUSTAT for the period between 1980 and 2015. Growth rates are computed controlling for aggregate fluctuations, heterogeneity, and relative size. We find that all variables' growth rates are largely leptokurtic, with kurtosis ranging between 6 and 11 (the benchmark kurtosis is 3 for a Normal random variable). We interpret this evidence as suggesting the effect of large infrequent shocks, which generate leptokurtic distributions of firm outcomes beyond prices.

Signals Firms do not observe their markup gaps directly. They receive continuous noisy observations about the markup gap, denoted by s_t, which evolve according to

$$ds_t = \mu_t dt + \gamma dZ_t \tag{2}$$

where the signal noise Z_t follows a Wiener process, independent from W_t. The volatility parameter γ measures the information friction's size. Note that the underlying state, μ_t, enters as the drift of the signal. This representation makes the filtering problem tractable as the signal has continuous paths.[5] This signal extraction problem can be reinterpreted, when written in discrete time, as a problem with undistinguishable permanent and transitory shocks. The signal noise can be reinterpreted as transitory volatility affecting the state. This alternative interpretation is useful for building the economic interpretation of our model. See Section H in Online Appendix for details.

Information set We assume that a firm knows if there has been an infrequent large shock to her markup—our notion of a regime change—, but not the size of the innovation u_t. This assumption implies that the information set at time t is given by the σ-algebra generated by the history of signals s and realizations of Q:

$$I_t = \sigma\{s_r, Q_r;\ r \leq t\}$$

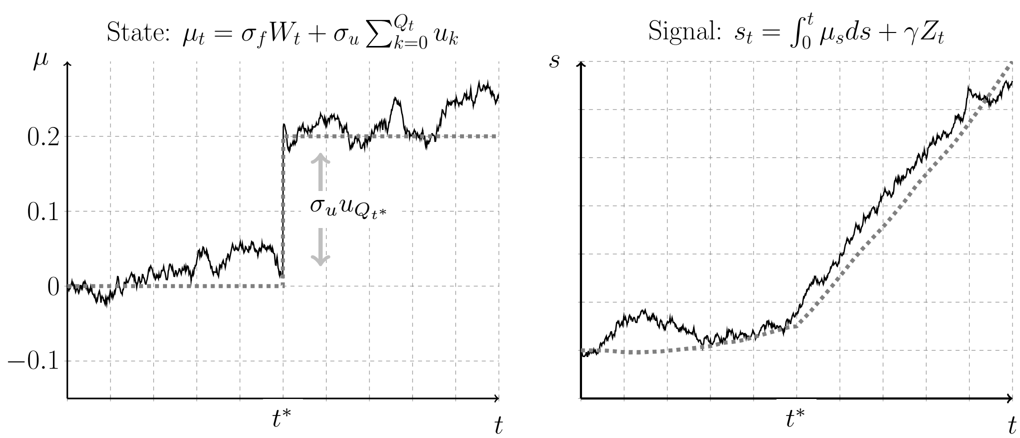
Since Poisson innovations are not observed but have conditional mean of zero, firms know that the arrival of a regime change could push her markups either upwards or downwards, but in expectation it would have no effect. Thus regime changes reflect innovations in the economic environment that, given the information available to her, a firm cannot assign a sign or magnitude to the effects it will have on her markup. We have set $\mathbb{E}[u_t] = 0$ for two reasons: it makes the algebra more tractable, and more importantly, it allows us to match the price change distribution's symmetry around zero. However, it is not a crucial assumption. All the results can be easily extended to the case of positive or negative conditional mean. Furthermore, we can also relax the assumption about u's observability and include partial information about the size or sign of its realization. For example, we could include an additional signal about u that the filtering would take into account when estimating the state. As long as the shock's timing is known, we can make different assumptions about u and maintain analytical traction.

[5]Rewrite the signal as $s_t = \int_0^t \mu_s ds + \gamma Z_t$ which is the sum of an integral and a Wiener process, and therefore it is continuous. See Chapter 6 in Øksendal (2007) and the Appendix for more details on filtering problems in continuous time.

The crucial assumption is that the firm knows the arrival of a regime change. This allows us to keep the problem within a finite dimensional state Gaussian framework, as we show in Proposition 1, where only the first two moments of posterior distributions are needed for the firm's decision problem. Another approach would be to assume a finite number of markup gaps and keep track of their probability distribution, and use the techniques of hidden state Markov models pioneered by Hamilton (1989). Other methods that would solve the filtering problem without our assumptions involve approximations as in the Kim (1994) filter or particle filters. These alternative methods have infinite or very large state spaces and the curse of dimensionality makes them unsuitable for solving the inaction problem.

Figure I illustrates the evolution of the markup gap and the signal process. It assumes that there is a regime change at time t^*. At that moment, the average *level* of the markup gap jumps to a new value; nevertheless, the signal has continuous paths and only its *slope* changes to a new average value.

Figure I – Illustration of the Markup Gap and the Signal Processes



Left panel: describes a sample path of the markup gap. The dashed line describes the compound Poisson process and the solid line describes the markup gap (the sum of the compound Poisson process and the Wiener process). t^* is the date of an increase in the Poisson counter. Right panel: describes a sample path for the signal. The dashed line describes the drift and the solid line describes the signal (the sum of the drift and the local volatility).

2.2 Filtering problem

This section describes the filtering problem and derives the laws of motion for estimates and estimation variance, our measure of uncertainty. The key challenge is to keep the finite state properties of the Gaussian model and apply Bayesian estimation in a jump-diffusion framework. Álvarez, Lippi and Paciello (2011) analyzes the filtering problem without the jumps and it shows that the steady state of such a model is equal to a perfect information model. Our contribution extends the Kalman–Bucy filter beyond the standard assumption of Brownian motion innovations. We are able to represent the posterior distribution of markup gaps $\mu_t|\mathcal{I}_t$ as a function of mean and variance. To our knowledge, this is a novel result in the filtering literature.

Firms make estimates in a Bayesian way by optimally weighing new information contained in signals against old information from previous estimates. This is a passive learning technology in the sense that firms process the information that is available to them, but they cannot make any action to change the quality of the signals; this contrasts with the active learning models in Keller and Rady (1999), Bachmann and Moscarini (2011), Willems (2013), and Argente and Yeh (2015) where firms learn the elasticity of their demand by experimenting with price changes.

Estimates and uncertainty Let $\hat{\mu}_t \equiv \mathbb{E}[\mu_t|I_t]$ be the best estimate (in a mean-squared error sense) of the markup gap and let $\Sigma_t \equiv \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|I_t]$ be its variance. Firm level uncertainty is defined as $\Omega_t \equiv \frac{\Sigma_t}{\gamma}$, which is the estimation variance normalized by the signal volatility. Proposition 1 below establishes the laws of motion for estimates and uncertainty for our drift-less case. In the Appendix we provide the generalization of the Kalman-Bucy filter to a jump-diffusion process with drift.

Proposition 1 (**Filtering equations**). *Let the markup gap and the signal evolve according to the following processes:*

$$\begin{aligned} (\textit{state}) \qquad d\mu_t &= \sigma_f dW_t + \sigma_u u_t dQ_t, \qquad && \mu_0 \sim \mathcal{N}(a,b) \\ (\textit{signal}) \qquad ds_t &= \mu_t dt + \gamma dZ_t, && s_0 = 0 \end{aligned}$$

where W_t, Z_t are Wiener processes, Q_t is a Poisson process with intensity λ, $u_t \sim \mathcal{N}(0,1)$, and a, b are constants. Let the information set be given by $\mathcal{I}_t = \sigma\{s_r, Q_r; r \leq t\}$, and define the markup estimate $\hat{\mu}_t \equiv \mathbb{E}[\mu_t|\mathcal{I}_t]$ and the estimation variance $\Sigma_t \equiv \mathbb{V}[\mu_t|\mathcal{I}_t] = \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|\mathcal{I}_t]$. Finally, define firm uncertainty as the estimation variance normalized by the signal noise: $\Omega_t = \frac{\Sigma_t}{\gamma}$. Then the posterior distribution of markups is Gaussian $\mu_t|\mathcal{I}_t \sim \mathcal{N}(\hat{\mu}_t, \gamma\Omega_t)$, where $(\hat{\mu}_t, \Omega_t)$ satisfy

$$d\hat{\mu}_t = \Omega_t d\hat{Z}_t, \qquad \hat{\mu}_0 = a \tag{3}$$

$$d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t, \qquad \Omega_0 = \frac{b}{\gamma} \tag{4}$$

$\hat{Z}_t$ is the innovation process given by $d\hat{Z}_t = \frac{1}{\gamma}(ds_t - \hat{\mu}_t dt) = \frac{1}{\gamma}(\mu_t - \hat{\mu}_t)dt + dZ_t$ and it is one-dimensional Wiener process under the probability distribution of the firm, and it is independent of dQ_t.

Proof. All proofs are given in the Appendix. □

The proof consists of three steps. First, we show that the solution to the system of stochastic differential equations in (1) and (2), conditional on the history of Poisson shocks, follows a Gaussian process; second, we show that $\mu_t|\mathcal{I}_t$ is a Gaussian random variable where its mean and variance can be obtained as the limit of a discrete sampling of observations; and third, we show that the laws of motion of markup estimates and uncertainty obtained with discrete sampling converge to the system given by (3) and (4). We now discuss each filtering equation with detail.

Higher uncertainty implies more volatile estimates Equation (3) says that the estimate $\hat{\mu}_t$ is a Brownian motion driven by the innovation process $\hat{Z}_t$ with stochastic volatility with jumps given by Ω_t. We can see this property using a discrete time approximation of the estimates process in (3) and the signal process in (2). Consider a small period of time Δ. The markup gap estimate at time $t + \Delta$ is given by the Bayesian convex combination of the previous estimate $\hat{\mu}_t$ and the signal change $s_t - s_{t-\Delta}$ (see Appendix for a formal proof)

$$\hat{\mu}_{t+\Delta}\Delta = \underbrace{\frac{\gamma}{\Omega_t\Delta + \gamma}}_{\text{weight on prior estimate}} \hat{\mu}_t\Delta + \underbrace{\left(1 - \frac{\gamma}{\Omega_t\Delta + \gamma}\right)}_{\text{weight on signal}} (s_t - s_{t-\Delta}) \tag{5}$$

A discrete time approximation of the signal is given by:

$$s_t = s_{t-\Delta} + \mu_t \Delta + \gamma \sqrt{\Delta} \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1) \tag{6}$$

Substituting (6) into (5) and rearranging we obtain:

$$\hat{\mu}_{t+\Delta} - \hat{\mu}_t = \frac{\Omega_t}{\Omega_t \Delta + \gamma} \underbrace{\Big((\mu_t - \hat{\mu}_t)\Delta + \gamma \sqrt{\Delta} \epsilon_t \Big)}_{\to\ \gamma d\hat{Z}_t} \tag{7}$$

Since the estimate $\hat{\mu}_t$ is unbiased, the term inside parentheses has all the properties of a Wiener process. Therefore, $\hat{\mu}_t$ follows an Itō process with local variance given by Ω_t. The approximation in (5) makes evident that, when uncertainty is high, the estimates put more weight on the signals than on the previous estimate. This means that the estimate incorporates more information about the current markup μ_t; in other words learning is faster, but it also brings more white noise ϵ_t into the estimation. Estimates become more volatile with high uncertainty. This effect will be key in our discussion of firms' responsiveness to monetary shocks, as with high uncertainty the markup estimates will incorporate the monetary shock faster and responsiveness will be larger.

Uncertainty cycles Equation (4) shows that uncertainty has a deterministic and a stochastic component, where the latter is active whenever the markup gap receives a regime change. Let us study each component separately. In the absence of regime changes ($\lambda = 0$), uncertainty Ω_t follows a deterministic path which converges to the constant volatility of the continuous shocks σ_f, i.e. the volatility of the true state. The deterministic convergence is a result of the learning process: as time goes by, estimation variance decreases until the only volatility left is that of the state. In the model with regime changes ($\lambda > 0$), uncertainty jumps up on impact with the arrival of regime change and then decreases deterministically until the arrival of a new regime change that will push uncertainty up again. The time series profile of uncertainty features a saw-toothed profile that never stabilizes due to the recurrent nature of these shocks. If the arrival of the infrequent shocks were not known and instead the firm had to filter their arrival as well, uncertainty would feature a hump-shaped profile instead of a jump. Although uncertainty never settles down, it is convenient to characterize the level of uncertainty such that its expected change is equal to zero, $\mathbb{E}\left[d\Omega_t \middle| \mathcal{I}_t\right] = 0$. It is equal to the variance of the state $\mathbb{V}[\mu_t] = {\Omega^*}^2 t$, hence we call this *fundamental uncertainty* with a value of $\Omega^* \equiv (\sigma_f^2 + \lambda \sigma_u^2)^{\frac{1}{2}}$. The ratio of current to fundamental uncertainty Ω_t/Ω^* appears in decision rules and price statistics.

Further comments on the filtering problem A notable characteristic of this filtering problem is that point estimates, as well as the signals and innovations, have continuous paths even though the underlying state is discontinuous. The continuity of these paths comes from two facts. First, changes in the state affect the slope of the innovations and signals but not their levels; second, the expected size of an infrequent shock u_t is zero. As a consequence of the continuity, markup estimations are not affected by the arrival of a regime change; only uncertainty features jumps. It is also worth noticing that both the filtered estimates $\mu_t|\mathcal{I}_t$ and smoothed estimates $\mu_{t-\delta}|\mathcal{I}_t$ with $\delta > 0$ are Gaussian. In contrast, the predicted estimate $(\mu_{t+\delta}|\mathcal{I}_t)$ is not. For instance, in the case $\sigma_f = 0$, the predicted markup converges to a Laplace distribution with fat tails. We focus our attention on the filtered estimate since it is the only input in our firm's decision problem. We leave for further research the analysis of other estimates.

2.3 Decision rules

With the filtering problem at hand, this section derives the price adjustment decision of the firm.

Sequential problem Let $\{\tau_i\}_{i=1}^{\infty}$ be the series of dates where the firm adjusts her markup gap and $\{\mu_{\tau_i}\}_{i=1}^{\infty}$ the series of reset markup gaps on the adjusting dates. Given an initial condition μ_0, the law of motion for markup gaps, and the filtration $\{\mathcal{I}_t\}_{t=0}^{\infty}$, the sequential problem of the firm is described by:

$$\max_{\{\mu_{\tau_i},\tau_i\}_{i=1}^{\infty}} -\mathbb{E}\left[\sum_{i=0}^{\infty} e^{-r\tau_{i+1}}\left(\theta+\int_{\tau_i}^{\tau_{i+1}} e^{-r(s-\tau_{i+1})}B\mu_s^2\,ds\right)\right] \tag{8}$$

The sequential problem is solved recursively as a stopping time problem using the Principle of Optimality (see Øksendal (2007) and Stokey (2009) for details). This is formalized in Proposition 2. The firm's state has two components: the point estimate of the markup gap $\hat{\mu}$ and the level of uncertainty Ω attached to that estimate. Given her current state $(\hat{\mu}_t, \Omega_t)$, the firm policy consists of (i) a stopping time τ, which is a measurable function with respect to the filtration $\{\mathcal{I}_t\}_{t=0}^{\infty}$; and (ii) the new markup gap μ'.

Proposition 2 **(Stopping time problem).** *Let $(\hat{\mu}_0, \Omega_0)$ be the firm's current state immediately after the last markup adjustment. Also let $\bar{\theta} = \frac{\theta}{B}$ be the normalized menu cost. Then the optimal stopping time and reset markup gap (τ, μ') solve the following problem:*

$$V(\hat{\mu}_0,\Omega_0) = \max_{\tau}\mathbb{E}\left[\int_0^{\tau} -e^{-rs}\hat{\mu}_s^2 ds + e^{-r\tau}\Big(-\bar{\theta}+\max_{\mu'} V(\mu',\Omega_\tau)\Big)\Big|\mathcal{I}_0\right] \tag{9}$$

subject to the filtering equations in Proposition 1.

Observe in Equation (9) that the estimates enter directly into the instantaneous return, while uncertainty affects only the continuation value. To be precise, uncertainty does have a negative effect on current profits that reflects the firm's permanent ignorance about her true productivity. However, this loss is constant and can be treated as a sunk cost; thus it is set to zero.

Inaction region The solution to the stopping time problem is characterized by an inaction region $\mathcal{R}$ such that the optimal time to adjust is given by the first time that the state falls outside such a region:

$$\tau = \inf\{t > 0 : (\mu_t, \Omega_t) \notin \mathcal{R}\}$$

Since the firm has two states, the inaction region is two-dimensional. Let $\bar{\mu}(\Omega)$ denote the inaction region's border as a function of uncertainty. The inaction or continuation region is described by the set:

$$\mathcal{R} = \{(\mu,\Omega) : |\mu| \leq \bar{\mu}(\Omega)\}$$

The symmetry of the inaction region around zero is inherited from the specification of the stochastic process, the quadratic profits, and zero inflation. Notice that this is a non-standard inaction problem since it is two-dimensional, and moreover, there is a jump process in the Ω dimension. In order to provide sufficient conditions of optimality, we impose the Hamilton-Jacobi-Bellman equation, the value matching condition, and, following Theorem 2.2 in Øksendal and Sulem (2010), we ensure that the standard smooth pasting condition is satisfied by both states.

Section B of the Online Appendix verifies that the conditions in that Theorem hold in our problem; and Section C.3 verifies numerically that the smooth pasting conditions for $\hat{\mu}$ and Ω are valid. Proposition 3 formalizes these points.

Proposition 3 (**HJB Equation, Value Matching and Smooth Pasting**). *Let $\phi : R \times R^+ \to R$ be a function and let ϕ_x denote the derivative of ϕ with respect to x. Assume ϕ satisfies the following conditions:*

1. *For all states in the interior of the inaction region $\mathcal{R}^o$, ϕ solves the Hamilton-Jacobi-Bellman (HJB) equation:*

$$r\phi(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 - \Omega^2}{\gamma}\right)\phi_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2}\phi_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda\left[\phi\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - \phi(\hat{\mu}, \Omega)\right] \tag{10}$$

2. *At the border of the inaction region $\partial\mathcal{R}$, ϕ satisfies the value matching condition, which sets the value of adjusting equal to the value of not adjusting:*

$$\phi(0, \Omega) - \bar{\theta} = \phi(\bar{\mu}(\Omega), \Omega) \tag{11}$$

3. *At the border of the inaction region $\partial\mathcal{R}$, ϕ satisfies two smooth pasting conditions, one for each state:*

$$\phi_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega) = 0, \quad \phi_\Omega(\bar{\mu}(\Omega), \Omega) = \phi_\Omega(0, \Omega) \tag{12}$$

Then ϕ is the value function $\phi = V$ and $\tau = \inf\{t > 0 : \phi(0, \Omega_t) - \theta > \phi(\hat{\mu}_t, \Omega_t)\}$ is the optimal stopping time.

A key property of the HJB is the lack of interaction terms between uncertainty and markup gap estimates. This property is implied by the passive learning process in which the firm cannot change the quality of the information flow by changing her markup. Using the HJB equation and other conditions, Proposition 4 gives an analytical characterization of the inaction region's border $\bar{\mu}(\Omega)$. The proof uses a Taylor expansion of the value function. Section C of the Online Appendix compares the approximation of the policy with its exact counterpart computed numerically and concludes that the approximation is adequate in the parameter space of interest. We do the same comparison for the conditional moments computed in the next sections.

Proposition 4 (**Inaction region**). *For r and $\bar{\theta}$ be small, the border of the inaction region is approximated by*

$$\bar{\mu}(\Omega) = \left(\frac{6\bar{\theta}\Omega^2}{1 + \mathcal{L}^{\bar{\mu}}(\Omega)}\right)^{1/4}, \quad \textit{with} \quad \mathcal{L}^{\bar{\mu}}(\Omega) = \left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{\Omega}{\Omega^*} - 1\right) \tag{13}$$

The elasticity of $\bar{\mu}(\Omega)$ with respect to Ω is equal to

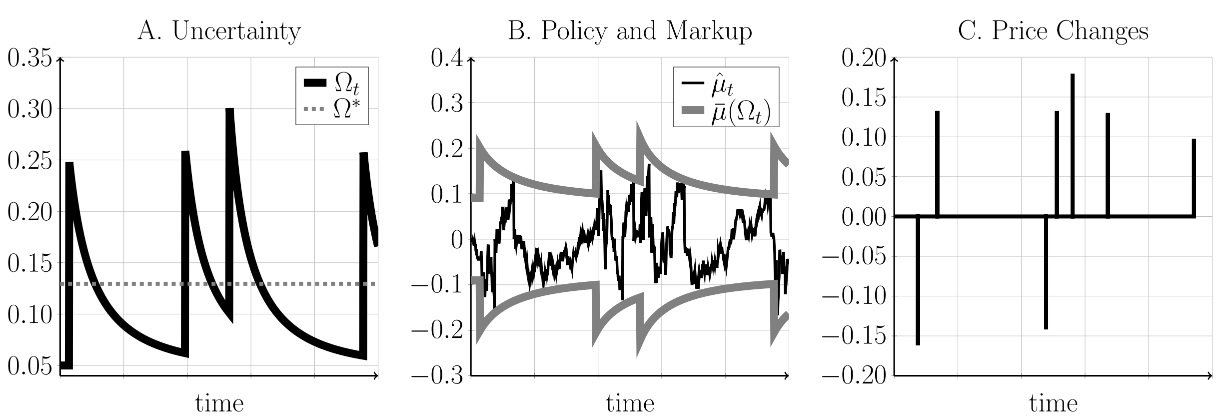
$$\mathcal{E}(\Omega) \equiv \frac{1}{2} - \left(\frac{1}{6}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\frac{\Omega}{\Omega^*} \tag{14}$$

Lastly, the reset markup gap is equal to $\hat{\mu}' = 0$.

Higher uncertainty implies wider inaction region The numerator of the inaction region $\bar{\mu}(\Omega)$ in equation (13) is increasing in uncertainty and captures the well known *option value effect* (see Barro (1972) and Dixit (1991)). As a result of belief dynamics, the option value is time varying and driven by uncertainty. In the denominator there is a new factor $\mathcal{L}^{\bar{\mu}}(\Omega)$ that amplifies or dampens the option value effect depending on the ratio of current uncertainty to fundamental uncertainty $\frac{\Omega}{\Omega^*}$. When current uncertainty is high with respect to its average level $\left(\frac{\Omega}{\Omega^*} > 1\right)$, uncertainty is expected to decrease ($\mathbb{E}[d\Omega] < 0$) and therefore future option values also decrease. This feeds back into the current inaction region shrinking it as $\mathcal{L}^{\bar{\mu}}(\Omega) > 0$. Analogously, when uncertainty is low with respect to its average level $\left(\frac{\Omega}{\Omega^*} < 1\right)$, it is expected to increase ($\mathbb{E}[d\Omega] > 0$) and thus the option values in the future also increase. This feeds back into current bands that get expanded as $\mathcal{L}^{\bar{\mu}}(\Omega) < 0$. The overall effect of uncertainty on the inaction region also depends on the ratio of the normalized menu cost and the signal noise. The expression (13) shows that small menu costs θ paired with large signal noise γ make the factor $\mathcal{L}^{\bar{\mu}}(\Omega)$ close to zero, implying that the elasticity of the inaction region with respect to uncertainty $\mathcal{E}(\Omega)$ in (14) is close to 1/2 and thus the inaction region is increasing in uncertainty. The critical result, which will be used later in characterizing micro price statistics, is that the elasticity of the inaction region to uncertainty is less than unity.

Figure II shows a particular firm realization for the parametrization we will use in our quantitative exercise, which has small menu costs $\bar{\theta}$ and large signal noise γ. Panel A shows the evolution of uncertainty, which follows a saw-toothed profile: it decreases monotonically with learning until a regime change happens and makes uncertainty jump up; then, learning brings uncertainty down again. The dashed horizontal line is fundamental uncertainty Ω^*. Panel B plots the markup gap estimate and the inaction region. The inaction region follows uncertainty's profile because the calibration makes the inaction region increasing in uncertainty. Finally, Panel C shows the magnitude of price changes. These changes are triggered when the markup gap estimate touches the border of the inaction region.

Figure II – Sample Paths For One Firm



Panel A: Uncertainty (solid line) and fundamental uncertainty (horizontal dotted line). Panel B: Markup gap estimate (solid line) and inaction region (dotted line). Panel C: Magnitude of price changes. This figure simulates one realization of the stochastic processes using the finite difference method described in Section C of the Online Appendix, and uses the analytical approximation of the inaction region to compute the policy and price changes.

Note that without regime changes, uncertainty would converge to a constant, i.e., $\Omega \to \sigma_f$. The inaction region would also become constant and akin to that of a steady state model without information frictions, namely $\bar{\mu} = \left(6\bar{\theta}\sigma_f{}^2\right)^{1/4}$. That is the case analyzed in the Online Appendix in Álvarez, Lippi and Paciello (2011). As that paper shows, such a model collapses to that of Golosov and Lucas (2007) where there is no price change size dispersion, since all firms would have the same inaction region. Therefore, both the regime changes and the information friction are key to generate the cross-sectional variation in price setting that arises from the heterogenous uncertainty.

How does uncertainty affect the adjustment frequency? Notice that price changes appear to be clustered over time, that is, there are recurrent periods with high adjustment frequency followed by periods of low adjustment frequency. Figure II shows that after a regime change arrives, the estimation becomes more volatile, which increases the probability of hitting the bands and changing the price. As a response to higher volatility and to save on menu costs, the inaction region becomes wider, which reduces the probability of a price change. Therefore, we have two opposite forces acting on the adjustment frequency. Since the elasticity of the inaction region with respect to uncertainty is less than unity, the volatility effect dominates and higher uncertainty brings more price changes. We formalize these observations in the following section on price statistics.

3 Uncertainty and micro price statistics

In this section we characterize analytically two price statistics that are crucial to understand the economy's response to aggregate nominal shocks: the expected duration of prices and the hazard rate of price adjustment. First, we focus on price statistics *conditional* on a level of uncertainty, and we shed light on the role of uncertainty in pricing behavior. We show that higher uncertainty decreases price duration (increases the adjustment frequency) and that the hazard rate of price adjustment is decreasing for firms with a high level of uncertainty. Furthermore, we show that the hazard rate's slope is determined by the volatility of the signal noise. To obtain these results, we require an elasticity of the inaction region with respect to uncertainty that is less than unity.

Second, we aggregate the conditional statistics to generate the *unconditional* statistics that we observe in the data. For aggregation, we use the renewal distribution of uncertainty, which is the distribution of uncertainty of adjusting firms. We show that this renewal distribution puts more weight on high levels of uncertainty than does the steady state distribution of uncertainty. This implies that aggregate statistics reflect the behavior of highly uncertain firms, and therefore, decreasing hazard rates are also observed in the aggregate.

3.1 Expected time

In Proposition 5 we establish a positive relationship between adjustment frequency and uncertainty, as observed in Figure II. It is followed by Proposition 6 which formalizes a positive relationship between adjustment frequency and uncertainty dispersion. These relationships prove to be very useful to back out an unobservable state—uncertainty—with observable price statistics.

Proposition 5 (Conditional Expected Time). *Let r and $\bar{\theta}$ be small. The expected time for the next price change conditional on the state, denoted by $\mathbb{E}[\tau\big|\hat{\mu},\Omega]$, is approximated as:*

$$\mathbb{E}[\tau\big|\hat{\mu},\Omega] = \frac{\bar{\mu}(\Omega)^2-\hat{\mu}^2}{\Omega^2}\left(1+\mathcal{L}^\tau(\Omega)\right) \quad \textit{where} \quad \mathcal{L}^\tau(\Omega)\equiv 2\left(\frac{\Omega}{\Omega^*}-1\right)\left(1-\mathcal{E}(\Omega^*)\right)\left(\frac{\gamma(24\bar{\theta})^{1/2}}{\gamma+(24\bar{\theta})^{1/2}}\right) \tag{15}$$

If the elasticity of the inaction region with respect to uncertainty is lower than unity and signal noise is large, then the expected time between price changes (i.e. $\mathbb{E}[\tau\big|0,\Omega]$) is a decreasing and convex function of uncertainty.

The expected time between price changes has two terms. The first term $\frac{\bar{\mu}(\Omega)^2-\hat{\mu}^2}{\Omega^2}$ is standard, and it states that the closer the current markup gap is to the border of the inaction region, then the shorter the expected time for the next adjustment. This term is decreasing in uncertainty with an elasticity larger than unity in absolute value, and it is time varying. The second term $\mathcal{L}^\tau(\Omega)$ amplifies or dampens the first effect depending on the level of uncertainty, and it has an elasticity equal to unity with respect to uncertainty. Therefore, uncertainty's overall effect on the expected time to adjustment is negative: a high uncertainty adjusts more frequently than a low uncertainty firm.

Notice that if menu costs are small and signal noise is large, the expected time between price changes is given by the ratio of the inaction region to uncertainty: $\mathbb{E}[\tau|0,\Omega] = \left(\frac{\bar{\mu}(\Omega)}{\Omega}\right)^2$. Since that the elasticity of the inaction region with respect to uncertainty is less than unity, the expected time decreases with uncertainty. There is empirical evidence of this relationship. Bachmann, Born, Elstner and Grimme (2013) use German survey data to document a positive relationship between firm-level belief uncertainty, measured as the variance of sales' forecast errors, and the individual adjustment frequency; Vavra (2014) uses BLS micro price data to document a positive relationship between the cross-sectional dispersion of price changes—another measure of uncertainty—and the individual frequency of price changes.

The next result generalizes Proposition 1 in Álvarez, Le Bihan and Lippi (2014) for the case of heterogeneous uncertainty. It establishes a positive relationship between uncertainty dispersion and adjustment frequency, and between uncertainty dispersion and price change dispersion. Blanco (2016a) shows that a similar result holds in the case of positive inflation.

Proposition 6 (**Uncertainty and Frequency**). *The following relationship between uncertainty dispersion, average price duration, and price change dispersion holds:*

$$\mathbb{E}[\Omega^2] = \frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]} \tag{16}$$

Holding fixed uncertainty's cross-sectional dispersion in the left-hand side, expression (16) establishes a positive link between average price duration and price change dispersion. Prices either change often for small amounts or rarely for large amounts. This implication of menu cost models can be tested empirically, for instance, using price statistics from different sectors. As an alternative way to read this relationship, consider a fixed price change dispersion; then heterogeneity in uncertainty and average price duration are negatively related. Underlying these results is a Jensen inequality and the fact that frequency decreases with price age.

The key point of the previous proposition is that observable price statistics provide a way to recover statistical moments of an unobserved state. For empirical applications, Proposition 6 can be applied to micro data to recover a measure of firm level uncertainty.

3.2 Hazard rate

We turn next into characterizing the hazard rate, which is a dynamic measure of adjustment frequency. Let $h_\tau(\Omega)$ be the conditional hazard rate of price adjustment. It is the probability of changing the price at date τ since the last price change, and it is conditional on a current level of uncertainty Ω. It is computed as $h_\tau(\Omega) \equiv \frac{f(\tau|\Omega)}{\int_\tau^\infty f(s|\Omega)ds}$, where $f(s|\Omega)$ is the conditional distribution of stopping times. It reflects the probability of exiting the inaction region, or first passage time. Without loss of generality, assume the last adjustment occurred at time $t = 0$ and denote price duration with $\tau > 0$. The hazard rate is a function of two objects:

i) estimate's unconditional variance: this is the variance of the estimate at a future date τ from a time $t = 0$ perspective, which we denote by $\mathcal{V}_\tau(\Omega_0)$, i.e. $\hat{\mu}_\tau|\mathcal{I}_0 \sim \mathcal{N}(0, \mathcal{V}_\tau(\Omega_0))$

ii) expected path of the inaction region $\bar{\mu}(\Omega)$ given the information available at time $t = 0$.

An analytical characterization of the hazard rate, conditional on an initial level of uncertainty Ω_0, is provided in Proposition 7. We make two assumptions, and their validity is tested in Section C.6 Online Appendix where we compute the exact numerical hazard rate. First, we assume that the inaction region is constant. This assumption is justified since our calibration implies a very small elasticity of the inaction region with respect to uncertainty. Second, we assume that after the last adjustment the firm expects no more Poisson shocks, which means that uncertainty will follow its deterministic path towards the volatility of the Brownian motion σ_f. Clearly, without the Poisson shocks, in steady state we would not have an initial level of uncertainty Ω_0 that is different from σ_f, but we can still think of the evolution of uncertainty given an initial condition. As our numerical results show, adding back the Poisson shocks does not produce significantly different hazards. The key message of the Proposition is that the concavity of the unconditional variance $\mathcal{V}_\tau(\Omega_0)$ determines the shape of the hazard function, because it measures how fast learning occurs, and the concavity is increasing in the initial level of uncertainty.

Proposition 7 (**Conditional Hazard Rate**). *Without loss of generality, assume the last price change occurred at $t = 0$ and let $\Omega_0 > \sigma_f$ be the initial level of uncertainty. The inaction region is constant $\bar{\mu}(\Omega_\tau) = \bar{\mu}_0$ and there are no infrequent shocks ($\lambda = 0$). Denote derivatives with respect to τ with a prime ($h'_\tau \equiv \partial h/\partial\tau$).*

1. *The estimate's unconditional variance, denoted by $\mathcal{V}_\tau(\Omega_0)$, is given by:*

$$\mathcal{V}_\tau(\Omega_0) = \sigma_f^2\tau + \mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) \tag{17}$$

where $\mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) \equiv \gamma(\Omega_0 - \Omega_\tau)$, with $\mathcal{L}_0^{\mathcal{V}}(\Omega_0) = 0$, $\lim_{\tau\to\infty} \mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) = \gamma(\Omega_0 - \sigma_f)$, and it is equal to:

$$\mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) = \gamma\Omega_0 - \gamma\sigma_f \left(\frac{\frac{\Omega_0}{\sigma_f} + \tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)} \right)$$

2. *$\mathcal{V}_\tau(\Omega_0)$ is increasing and concave in duration τ: $\mathcal{V}'_\tau(\Omega_0) > 0$ and $\mathcal{V}''_\tau(\Omega_0) < 0$. Furthermore, the following cross derivatives with initial uncertainty are positive:*

$$\frac{\partial \mathcal{V}_\tau(\Omega_0)}{\partial \Omega_0} > 0, \qquad \frac{\partial \mathcal{V}'_\tau(\Omega_0)}{\partial \Omega_0} > 0, \qquad \frac{\partial |\mathcal{V}''_\tau(\Omega_0)|}{\partial \Omega_0} > 0$$

3. *The hazard of adjusting the price at date τ, conditional on Ω_0, is characterized by:*

$$h_\tau(\Omega_0) = \frac{\pi^2}{8} \underbrace{\frac{\mathcal{V}'_\tau(\Omega_0)}{\bar{\mu}_0^2}}_{\text{decreasing in } \tau} \underbrace{\Psi\left(\frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)}_{\text{increasing in } \tau} \tag{18}$$

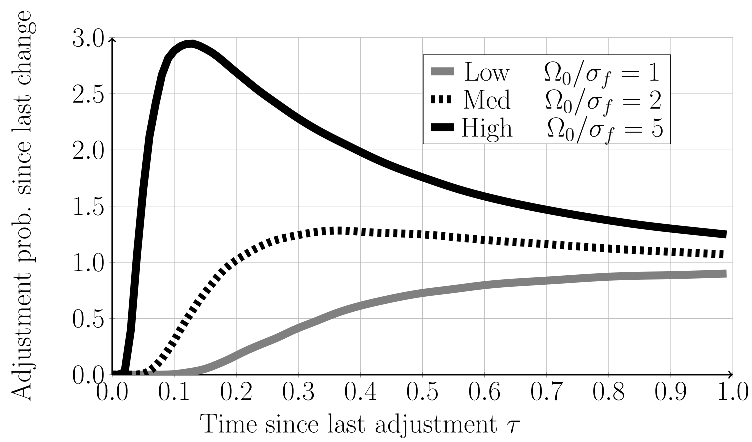
where $\Psi(x) \geq 0$, $\Psi(0) = 0$, $\Psi'(x) > 0$, $\lim_{x\to\infty} \Psi(x) = 1$, first convex then concave, and it is given by:

$$\Psi(x) = \frac{\sum_{j=0}^{\infty} \alpha_j \exp\left(-\beta_j x\right)}{\sum_{j=0}^{\infty} \frac{1}{\alpha_j} \exp\left(-\beta_j x\right)}, \quad \alpha_j \equiv (-1)^j(2j+1), \quad \beta_j \equiv \frac{\pi^2}{8}(2j+1)^2$$

4. *There exists a date $\tau^*(\Omega_0)$ such that the slope of the hazard rate is negative for $\tau > \tau^*(\Omega_0)$; and $\tau^*(\Omega_0)$ is decreasing in Ω_0.*

Estimate's unconditional variance $\mathcal{V}_\tau(\Omega_0)$ in (17) captures the evolution of uncertainty. The first term, $\sigma_f^2\tau$, refers to the linear time trend that comes from the fact that fundamental shocks follow a Brownian Motion. The second term, $\mathcal{L}_\tau^{\mathcal{V}}(\Omega_0)$, is an additional source of variance coming from imperfect information. The second point in Proposition 7 establishes that higher initial uncertainty increases the level, slope, and concavity of this additional variance. In other words, higher initial uncertainty brings higher expected gains from learning. In the third point, equation (18) shows that the imperfect information hazard rate given by the product of $\Psi(\cdot)$, an increasing function of τ, times the derivative of the unconditional variance $\mathcal{V}'_\tau$, a decreasing function of τ. The function $\Psi(\cdot)$ characterizes the hazard rate with perfect information which uses a transformation of the stopping time density by Kolkiewicz (2002). Therefore, there are two opposing forces acting upon the slope of the hazard rate and the hazard rate is non-monotonic. Finally, the fourth point states that there exists a date after which the hazard is downward sloping, and this date is shorter the higher initial uncertainty. Figure III illustrates the hazard rate for different initial conditions Ω_0. If the initial uncertainty is larger with respect to its lower bound σ_f, then the decreasing force becomes stronger and the hazard's slope is negative for a larger range of price durations.

Figure III – Hazard Rate Conditional on Initial Uncertainty



Conditional hazard rate of price adjustment for three different levels of initial uncertainty Ω_0, expressed as multiples of σ_f. These are approximated hazard rates with constant inaction regions and without Poisson shocks after the last adjustment. We use a larger σ_f than in the final calibration for illustration purposes.

Hazard rate and noise volatility The economics behind the non-monotonic hazard rate are as follows. Its increasing segment close to zero resembles the hazard rate of standard menu cost models, where the probability of an additional adjustment right after a price change is very low since the state has been reset and it lies in the middle of the inaction region. When initial uncertainty is very close to its minimum σ_f, there is no additional uncertainty and the hazard rate behaves as in a perfect information case. When initial uncertainty is very high, firms expects to transition from high uncertainty and frequent adjustments to low uncertainty and infrequent adjustments; this gives rise to the decreasing part. Then, the speed of the transition from high to low uncertainty is determined by the magnitude of information frictions, as captured by the noise volatility γ. If noise volatility is high, a firm will take a long time after a regime switch to learn her new level of permanent productivity. Both uncertainty and adjustment frequency remain high for many periods and the hazard rate is flat; in contrast, when the noise volatility is low, a firm learns quickly her new level of permanent productivity, both uncertainty and adjustment frequency fall after a few periods, and the hazard rate is relatively steep. This relationship between γ and the slope of the hazard rate will be exploited for the calibration of the model.

3.3 Aggregation

In the data we observe unconditional statistics. These moments are equal to the weighted average of the conditional statistics, where the weights are given by the renewal distribution of uncertainty. The renewal distribution is the stationary distribution of uncertainty *conditional* on price adjustment: it is the uncertainty faced by adjusting firms. Such distribution is different from the *unconditional* steady state distribution of uncertainty, which is the uncertainty in the entire cross-section. Importantly, micro price statistics are the outcomes of aggregation using the renewal distribution of uncertainty.

The distribution of price adjuster uncertainty—the renewal distribution—is difficult to compute analytically because of the jump process. Nevertheless, we can characterize the ratio between the renewal distribution and marginal distribution over uncertainty to show that it is increasing in uncertainty. The next proposition formalizes this result.

Proposition 8 (**Renewal distribution**). *Let $f(\hat{\mu},\Omega)$ be the joint density of markup gaps and uncertainty in the population of firms. Let $r(\Omega)$ be denote the density of uncertainty conditional on adjusting, or renewal density. Assume the inaction region is increasing in uncertainty (i.e. $\bar{\mu}'(\Omega) > 0$). Then we have the following results:*

1. *For each $(\hat{\mu},\Omega)$, we can write the joint density as $f(\hat{\mu},\Omega) = h(\Omega)g(\hat{\mu},\Omega)$, where $g(\hat{\mu},\Omega)$ is the density of markup gap estimates conditional on uncertainty and $h(\Omega)$ is the marginal density of uncertainty.*

2. *The ratio between the renewal and marginal densities of uncertainty is approximated by*

$$\frac{r(\Omega)}{h(\Omega)} \propto |g_{\hat{\mu}}(\bar{\mu}(\Omega),\Omega)|\Omega^2 \tag{19}$$

   where $g(\mu,\Omega)$ solves the following differential equation $\frac{\Omega^2-\Omega^{*2}}{\gamma} g_{\Omega}(\hat{\mu},\Omega) + \frac{\Omega^2}{2} g_{\hat{\mu}^2}(\hat{\mu},\Omega) = 0$ with border conditions: $g(\bar{\mu}(\Omega),\Omega) = 0$ and $\int_{-\bar{\mu}(\Omega)}^{\bar{\mu}(\Omega)} g(\mu,\Omega)d\mu = 1$.

3. *If $\Omega = \Omega^*$, then the ratio is proportional to the inverse of the expected time between price adjustments. Then if the inaction region's elasticity to uncertainty is lower than unity, the ratio is an increasing function of uncertainty:*

$$\frac{r(\Omega^*)}{h(\Omega^*)} \propto \frac{\Omega^{*2}}{2\bar{\mu}(\Omega^*)^2} = \frac{1}{2\mathbb{E}[\tau|(0,\Omega^*)]} \tag{20}$$

The key result of Proposition 8 is its last point, as it establishes that there is a greater mass of adjusters at high levels of uncertainty. Therefore, micro price statistics reflect more intensively the pricing behavior of highly uncertain firms. In the particular case of the hazard rate, the average hazard rate is decreasing because the renewal distribution puts a higher weight on the decreasing hazard rate of high uncertainty firms compared to the increasing hazard rate of low uncertainty firms. As before, this result is a direct consequence of having an elasticity of the inaction region to uncertainty lower than unity.

Belief uncertainty vs. stochastic volatility Uncertainty in this paper concerns idiosyncratic beliefs; it is the conditional variance of markup gap estimates. The volatility of the state is a known constant Ω^*; it is the realizations which are unknown. Our uncertainty shocks contrast with the stochastic volatility processes for productivity used in Vavra (2014) and Karadi and Reiff (2014). In these other papers, there is perfect information but the volatility of the state is stochastic. Regardless of the structure, however, the positive relationship between the frequency of price changes and the uncertainty (or volatility) faced by the firm is maintained.

Can we distinguish our model of imperfect information and endogenous uncertainty with one of perfect information and exogenous stochastic volatility, if the processes for uncertainty/volatility are the same? The answer is negative, as they are observationally equivalent. Can distinguish the models if these processes are different? The answer is positive. For instance, an autoregressive process for either stochastic uncertainty or volatility generates an increasing hazard rate, while the stochastic process with jumps generates a decreasing hazard rate. We compare such processes in the Online Appendix, Section D. The autoregressive process implies smooth changes in volatility; it generates an increasing hazard and a price change distribution with little dispersion and kurtosis compared to the Poisson model.

4 General equilibrium model

In this section we develop a standard general equilibrium framework with monopolistic firms that face the pricing problem with menu costs and information frictions studied in the previous sections. We will use this model to study the role of firm idiosyncratic uncertainty in the propagation of monetary shocks. For this purpose, we extend the environment in Golosov and Lucas (2007) to include the information friction and then characterize the steady state of the economy. We calibrate our economy to match several micro price statistics from CPI data in the United Kingdom computed. In particular, we calibrate the signal noise to match the slope of the hazard rate in the data, which we compute with a new methodology that eliminates survivor bias in its estimation.

4.1 Model

Environment Time is continuous. There is a representative consumer, a continuum of monopolistic firms, and a monetary authority.

Representative Household The household has preferences over consumption C_t, labor N_t, and real money holdings $\frac{M_t}{P_t}$, where P_t is the aggregate price level. She discounts the future at rate $r > 0$.

$$\mathbb{E}_0 \left[\int_0^\infty e^{-rt} \Big(\log C_t - N_t + \log \frac{M_t}{P_t} \Big) dt \right] \tag{21}$$

Consumption consists of a continuum of imperfectly substitutable goods indexed by z bundled together with a CES aggregator as

$$C_t = \Big(\int_0^1 \Big(A_t(z) c_t(z) \Big)^{\frac{\eta-1}{\eta}} dz \Big)^{\frac{\eta}{\eta-1}} \tag{22}$$

where $\eta > 1$ is the elasticity of substitution across goods and $c_t(z)$ is the amount of goods purchased from firm z at price $p_t(z)$. The ideal price index is the minimum expenditure necessary to deliver one unit of the final consumption good, and is given by:

$$P_t \equiv \left[\int_0^1 \left(\frac{p_t(z)}{A_t(z)} \right)^{1-\eta} dz \right]^{\frac{1}{1-\eta}} \tag{23}$$

In the consumption bundle and the price index, $A_t(z)$ reflects the quality of the good, with higher quality providing larger marginal utility of consumption but at a higher price. Quality shocks are firm specific and will be described fully in the firm's problem below. The household has access to complete financial markets. The budget includes income from wages W_t, profits Π_t from the ownership of all firms, and the opportunity cost of holding cash $R_t M_t$, where R_t is the nominal interest rate.

Let Q_t be the stochastic discount factor, or valuation in nominal terms of one unit of consumption in period t. Thus the budget constraint reads:

$$\mathbb{E}_0 \left[\int_0^\infty Q_t \left(P_t C_t + R_t M_t - W_t N_t - \Pi_t \right) dt \right] \leq M_0 \tag{24}$$

The household problem is to choose consumption of the different goods, labor supply and money holdings to maximize preferences (21) subject to (22), (23) and (24).

Monopolistic Firms On the production side, there is a continuum of firms indexed by $z \in [0,1]$. Each firm produces and sells her product in a monopolistically competitive market. They own a linear technology that uses labor as its only input: producing $y_t(z)$ units of good z requires $l_t(z) = y_t(z)A_t(z)$ units of labor, so that the marginal nominal cost is $A_t(z)W_t$ (higher quality $A_t(z)$ requires more labor input). For tractability, we assume that the quality shock enters both the production function and the marginal utility of the household, because this assumption helps to condense the numbers of states of the firm into one, the markup, as in Woodford (2009). Each firm sets a nominal price $p_t(z)$ and satisfies all demand at this posted price. Given the current price $p_t(z)$, the consumer's demand $c_t(z)$, and current quality $A_t(z)$, the instantaneous nominal profits of firm z are equal to the difference between nominal revenues and nominal costs:

$$\Pi(p_t(z), A_t(z)) = c_t(p_t(z), A_t(z)) \Big(p_t(z) - A_t(z) W_t \Big) \tag{25}$$

Firms maximize their expected stream of profits, which is discounted at the same rate of the consumer Q_t. They choose either to keep the current price or to change it, in which case they must pay a menu cost θ and reset the price to a new optimal one. Let $\{\tau_i(z)\}_{i=1}^\infty$ be a series of stopping times, that is, dates where firm z adjusts her price. The sequential problem of firm z is given by:

$$V(p_0(z), A_0(z)) = \max_{\{p_{\tau_i}(z), \tau_i(z)\}_{i=1}^\infty} \mathbb{E} \left[\sum_{i=0}^\infty Q_{\tau_{i+1}(z)} \left(-\theta + \int_{\tau_i(z)}^{\tau_{i+1}(z)} \frac{Q_s}{Q_{\tau_{i+1}(z)}} \Pi(p_{\tau_i(z)}, A_s(z)) ds \right) \right] \tag{26}$$

with initial conditions $(p_0(z), A_0(z))$ and subject to the quality process described next.

Quality process Firm z's log quality $a_t(z) \equiv \ln A_t(z)$ evolves as the following jump-diffusion process which is idiosyncratic and independent across z:

$$da_t(z) = \sigma_f W_t(z) + \sigma_u u_t(z) dQ_t(z) \tag{27}$$

where $W_t(z)$ is a Wiener process and $u_t(z)Q_t(z)$ is a compound Poisson process with arrival rate λ and Gaussian innovations $u_t(z) \sim \mathcal{N}(0,1)$ as in the previous sections. As before, firms do not observe their quality directly, and they do not learn it from observing their wage bill or revenues either. The only source of information are noisy signals $s_t(z)$ about quality together with the information that a regime change has hit them. The noisy signals $s_t(z)$ evolve as

$$ds_t(z) = a_t(z) dt + \gamma dZ_t(z) \tag{28}$$

where $Z_t(z)$ is an independent Brownian motion for each firm z and γ is signal noise. Each information set is $\mathcal{I}_t(z) = \sigma\{s_r(z), Q_r(z); r \leq t\}$. The parameters $\{\sigma_f, \sigma_u, \lambda, \gamma\}$ are identical across firms.

Money supply The monetary authority keeps money supply constant at a level $\bar{M}$.

4.2 Characterization of steady state equilibrium

Equilibrium A steady state equilibrium is a set of stochastic processes for (i) consumption strategies $c_t(z)$, labor supply N_t, and money holdings M_t for the household, (ii) pricing functions $p_t(z)$, (iii) prices W_t, R_t, Q_t, P_t, and (iv) a fixed distribution over firms F such that the household and firms optimize, markets clear at each date, and the distribution is consistent with actions.

Household optimality The first order conditions of the household problem establish: nominal wages as a proportion of the (constant) money stock $W_t = r\bar{M}$; the stochastic discount factor as $Q_t = e^{-rt}$; and demand for good z as $c_t(z) = A_t(z)^{\eta-1}\left(\frac{p_t(z)}{P_t}\right)^{-\eta} C_t$.

Constant aggregate prices The equilibrium with constant money supply implies a constant nominal wage $W_t = W$ and a constant nominal interest rate equal to the household's discount factor $R_t = 1 + r$. The ideal price index in (23) is also a constant $P_t = P$. Then nominal expenditure is also constant $P_tC_t = PC = M = W$. Therefore, there is no uncertainty in aggregate variables.

Back to quadratic losses Given the strategy of the consumer $c_t(z)$ and defining markups as $\mu_t(z) \equiv \frac{p_t(z)}{A_t(z)W}$, the instantaneous profits can be written as a function of markups alone:

$$\Pi(p_t(z), A_t(z)) \quad = \quad K\mu_t(z)^{-\eta}\Big(\mu_t(z) - 1\Big)$$

where $K \equiv M\left(\frac{W}{P}\right)^{1-\eta}$ is a constant in steady state. A second order approximation to this expression produces a quadratic form in the markup gap, defined as $\mu_t(z) \equiv \log(\mu_t(z)/\mu^*)$, i.e. the log deviations of the current markup to the unconstrained markup $\mu^* \equiv \frac{\eta}{\eta-1}$:

$$\Pi(\mu_t(z)) = C - B\mu_t(z)^2$$

where the constants are $C \equiv K\eta^{-\eta}(\eta-1)^{\eta-1}$ and $B \equiv \frac{1}{2}K\frac{(\eta-1)^{\eta}}{\eta^{\eta-1}}$. The constant C does not affect the decisions of the firm and it is omitted for the calculations of decision rules; the constant B captures the curvature of the original profit function. The firm's quadratic problem is the same as in Equation (8).

Markup gap estimation and uncertainty The markup gap is equal to[6]

$$\mu_t(z) = \log p_t(z) - a_t(z) - \log W - \log \mu^*$$

When the price is kept fixed (inside the inaction region), the markup gap is driven completely by the quality process: $d\mu_t(z) = -da_t(z)$. When there is a price adjustment, the markup process is reset to its new optimal value and then it will again follow the quality process. By symmetry of the Brownian motion without drift and the mean zero innovations of the Poisson process, we have that $da_t(z) = -da_t(z)$. Given the quality and signal processes in (27) and (28), together with $d\mu_t(z) = da_t(z)$, we obtain the same filtering equations as in Proposition 1, but now each process is indexed by z and is *iid* across firms:

$$\begin{aligned} d\hat{\mu}_t(z) &= \Omega_t(z)d\hat{Z}_t(z), \qquad \hat{Z}_t(z) \sim Wiener \\ d\Omega_t(z) &= \frac{\sigma_f^2 - \Omega_t^2(z)}{\gamma}dt + \frac{\sigma_u^2}{\gamma}dQ_t(z) \end{aligned}$$

[6]This expression shows that, under the Dixit-Stiglitz demand structure, quality $a_t(z)$ and optimal markup μ^* enter markup gaps in identical ways. Therefore, we could introduce fluctuations in demand elasticity instead of fluctuations in quality without changing the markup gap process.

Solution method The model is solved numerically as a discrete time version of the continuous time model. We approximate a discrete version of the firm value function with splines and solve it with iterative and colocation methods. To find the steady state, we compute the transition probability over a grid of states and recover the ergodic distribution as the eigenvector with unit eigenvalue. See Appendix for details.

4.3 Data and calibration

In this section we compute micro price statistics using disaggregated CPI data from the UK that will serve as calibration targets. The statistics are consistent with Dominick's database in Midrigan (2011) and the BLS monthly data in Nakamura and Steinsson (2008). For the estimation of the hazard rate—the key target—we propose a new methodology that controls for heterogeneity in adjustment frequency and eliminates survivor bias. The method uses the *relative* stopping times distribution, which are the stopping times normalized by the average duration of an item's price.

Data description We use monthly price quotes collected by the Office for National Statistics to construct the UK Consumer Price Index (CPI). There are several advantages in using this dataset: it is representative of the whole economy, it is publicly available from 1996 to 2016, and micro price statistics are very similar to other low-inflation countries such as the US, Canada and the EU. In total, there are 31 million price quotes, classified by sector and class level. We apply several filters and procedures to the data to make it compatible with the model. Following Klenow and Kryvtsov (2008) and Nakamura and Steinsson (2008) we filter out discounts and sectoral heterogeneity; we complete price quotes for missing observations and out-of-season products with the last available price; and we drop product substitutions, outliers, and months with changes in the VAT tax rate. CPI weights at the item level are used to construct moments. Other papers that use this data are Chu, Huynh, Jacho-Chavez and Kryvtsov (2016) and Kryvtsov and Vincent (2016).

Relative stopping times and relative hazard Denote items with $i = 1, 2, ..., N$ and their CPI weight with ω_i. Letting τ_i be a time between price changes for item i, the *relative stopping time* is τ_i divided by the average duration of that item through its sample length: $\hat{\tau}_i = \frac{\tau_i}{\mathbb{E}[\tau_i]}$. The overall stopping time distribution is then given by $\hat{\tau} = \hat{\tau}_i$ with probability ω_i. When using it to compute the hazard rate, this distribution does not generate survivor bias as is it not affected by heterogeneity across items. Section E in the Online Appendix, shows theoretically that this is indeed the case in three benchmark pricing models: menu cost with small frequent shocks, menu cost with leptokurtic shocks (Calvo-type economy), and Taylor model. To compute the hazard rate we use the distribution of $\hat{\tau}_i$ instead of τ_i and obtain a non-monotonic hazard rate which is increasing for the first 2 months and then decreasing. The challenge of this method is to get good estimates of $\mathbb{E}[\tau_i]$, because it requires sufficiently long time series for each item. To circumvent this issue, we exploit the structure of the UK database that samples the same item at different shops and locations with the purpose of constructing specific item-level inflation. Under the assumption of a similar pricing model across these two dimensions, it is possible to obtain a large sample of stopping times for each item to compute $\mathbb{E}[\tau_i]$.

Targets and Calibration We target the average adjustment size $\mathbb{E}[|\Delta p_i|] = 0.11$, price change dispersion $\sigma[\Delta p_i] = 0.13$, the average frequency of adjustment of $fr(\Delta p) = 0.107$, equivalent to an average price duration of $\mathbb{E}[\tau]$=9.3 months, and the shape of the hazard rate. The calibration is set at weekly frequency, and the price statistics are aggregated to match monthly price statistics in the data. The discount factor is set to $\frac{1}{1+r} = 0.96^{1/52}$ to match an annual risk free rate of 4%; the normalized menu cost is set to $\bar{\theta} = 0.064$ so that the expected menu cost payments represent 0.5% of the average revenue following the empirical evidence in Zbaracki *et al.* (2004) and Levy *et al.* (1997); and the CES elasticity is set to $\eta = 6$ to match an average markup of 20%.

We consider three alternative parametrizations of the stochastic process, but all match the same average frequency of adjustment. Table I and Figure IV summarize the following information. The first calibration shuts down the information friction ($\gamma = 0$) and the regime changes ($\lambda = 0$), and the only parameter σ_f is set to match the adjustment frequency. We consider this a simple version of the model in Golosov and Lucas (2007). The second calibration shuts down the information frictions ($\gamma = 0$) and the frequent shocks ($\sigma_f = 0$), keeping the regime changes active. Its other parameters λ and σ_u match the frequency and the dispersion of price changes. This model is a version of Gertler and Leahy (2008). The third is the full model with information frictions that has an additional parameter to calibrate, the signal noise, which is set to match the decreasing hazard rate. Importantly, we are able to match the same average adjustment frequency with a λ that is 60% smaller than in the second model. The reason is that for each Poisson shock prices change more than once because of the decreasing hazard; this is key to get the higher persistence of the output response to monetary shocks, as we show in the next section. Finally, the volatility of the frequent shocks, σ_f, is set very close to zero to get some small price changes.

Table I – Model Parameters and Data Targets

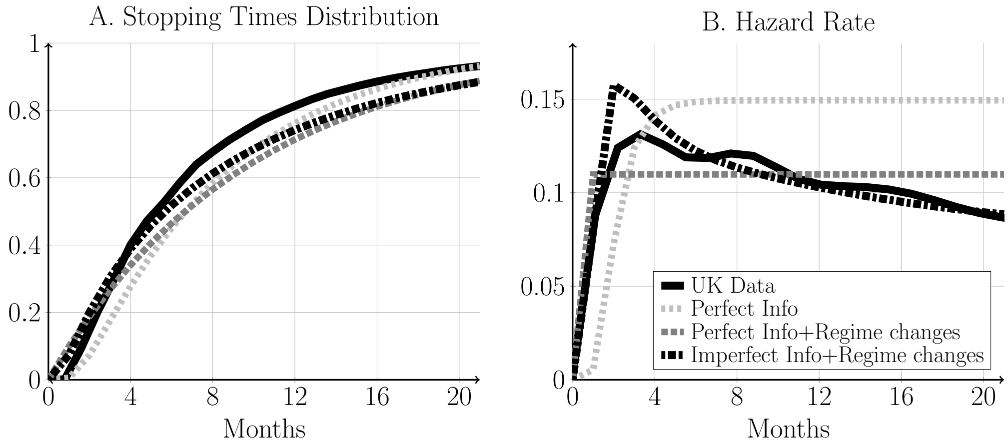
	UK Data	(1) Benchmark	(2) Regime Changes	(3) Info Frictions
Parameters				
σ_f		0.0188		0.0005
σ_u			0.10	0.17
λ			0.05	0.02
γ				0.37
Moments				
$\mathbb{E}[\tau]$ in months	9.31	9.31	9.50	9.51
$\mathbb{E}[\lvert\Delta p\rvert]$	0.11	0.10	0.12	0.13
$std[\Delta p]$	0.13	0.11	0.13	0.13
$kurtosis[\Delta p]$	3.95	1.04	1.61	1.63

Data: CPI Data from UK, 1996-2015. Models: (1) Perfect info with only frequent shocks; (2) Perfect info with only infrequent shocks; (3) Imperfect info with both types of shocks.

The imperfect information model obtains larger kurtosis than the other two, but still has some difficulty in matching the data, mainly because it has trouble generating small price changes as they are bounded below by the menu cost. In the Online Appendix, Section F, we extend the baseline model to the CalvoPlus model in Nakamura and Steinsson (2010), in which there are random opportunities to adjust prices without the menu cost. This extended model generates small price changes and a larger kurtosis of the price change distribution. Small price changes can also be generated by introducing economies of scope through multi-product firms as in Midrigan (2011) and Álvarez and Lippi (2014).

Figure IV shows the stopping time distribution and the hazard rate for the UK data and the three parametrizations of the model. The model with perfect information and only small shocks features an increasing hazard rate: after a price adjustment, it takes time for the small shocks to accumulate in the markup gap and trigger a price change. The model with perfect information and regime changes produces a flat hazard: the probability of changing the price is constant as it reflects the constant arrival rate of the Poisson shocks that trigger price changes. This result is at the core of Gertler and Leahy (2008) and Midrigan (2011) who show that a menu cost model with fat tailed shocks closely resembles a Calvo economy. Finally, the model with information frictions generates the decreasing hazard rate. Note that by calibrating one parameter, the signal noise γ, we can match very well the shape of the hazard rate for a large span of durations.

Figure IV – Stopping Times and Hazard Rates: Data and Models

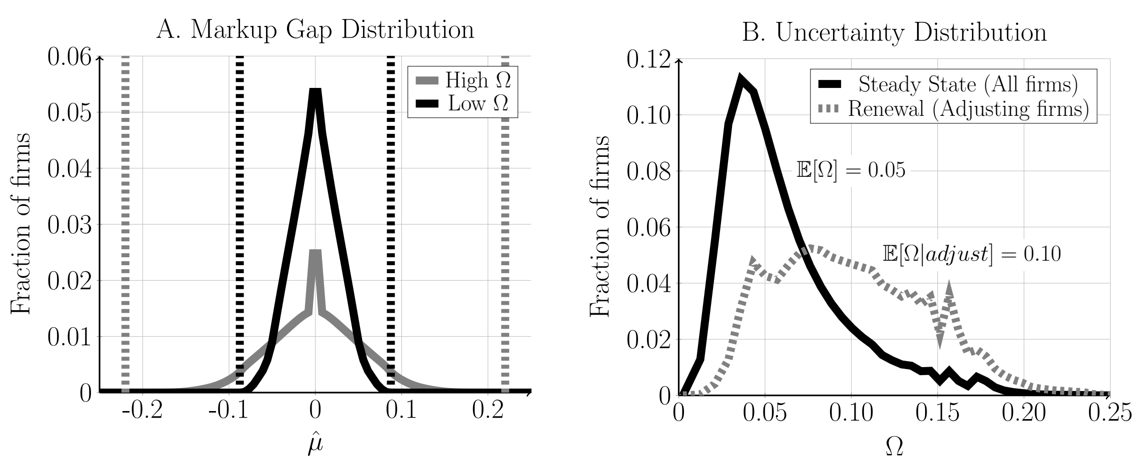


Distribution of stopping times and adjustment hazard rate for three parametrizations of the model and the data. Data: CPI Data from UK, 1996-2015. Models: (1) Perfect info with only frequent shocks; (2) Perfect info with only infrequent shocks; (3) Imperfect info with both types of shocks.

4.4 Steady state

Figure V shows the steady state distributions of markup gap estimates and uncertainty for the model with information frictions. Panel A plots the marginal distribution of markup gap estimates, $g(\hat{\mu}, \Omega)$, conditional on uncertainty being above or below its mean level. We observe that this distribution's support and dispersion are increasing in uncertainty. Average inaction region widths are $|\bar{\mu}(\Omega)| = 0.23$ and $|\bar{\mu}(\Omega)| = 0.08$ for high and low uncertainty firms, respectively. Panel B shows two uncertainty distributions. Consistent with Proposition 8, the steady state distribution of uncertainty $h(\Omega)$ is biased towards low uncertainty, and the expected level of uncertainty across all firms is equal to 0.05. In contrast, the renewal distribution $r(\Omega)$ shifts the mass towards higher uncertainty levels, and the expected level of uncertainty of adjusting firms is 0.1. As we said earlier, micro price statistics reflect the pricing behavior of highly uncertain firms.

Figure V – Steady State Distributions



Panel A: Steady state distribution of markup gap estimates, conditional on uncertainty. High uncertainty means uncertainty above mean, and low uncertainty below the mean. Panel B: Steady state and renewal distribution of uncertainty.

Uncertainty and age dependent statistics Our model generates a very tight connection between the age of price and firm uncertainty, where age is measured as the number of periods that a price has remained unchanged. High uncertainty firms are more likely to be charging young prices, while low uncertainty firms are more likely to be charging old prices. Therefore, price age becomes a determinant of the size and dispersion of price changes as well as the adjustment frequency. In particular, our model predicts that young (uncertain) prices are larger, more dispersed, and more likely to be reset than older (certain) prices.

These predictions are documented by Campbell and Eden (2014) using weekly scanner data. It defines a price as young if its age is less than three weeks and as old if its age is more than four weeks. That paper finds that conditional on adjustment, young prices have double the dispersion of old prices (15% vs. 7%) and that price changes in the extreme tails of the price change distribution tend to be young. Regarding the frequency, it finds that young prices are three times more likely to be changed than old prices (36% vs 13%). We compute analogous numbers in our model, defining young prices to be in the 25th quartile of the price age distribution and old prices to be in the 75th quartile. We obtain that dispersion of young price changes is one and half times larger than that of old prices, and that adjustment frequency is twice as large for young prices. Interestingly, the uncertainty faced by young prices is also twice the uncertainty faced by old prices; thus the relative adjustment frequency seems to be informative about the relative uncertainty faced by firms. Further evidence regarding age dependence in pricing is documented in Baley, Kochen and Sámano (2016). Using Mexican CPI data at the item-level, it shows that adjustment frequency and price change dispersion falls with the age of the price, as our model predicts.

5 Propagation of nominal shocks

What are the macroeconomic consequences of firm uncertainty? Specifically, how does output respond to an aggregate nominal shock in an economy where firms have heterogeneous uncertainty? In the first exercise, we compute the response of output to an unanticipated permanent monetary shock. We find that uncertainty heterogeneity amplifies the persistence of output response compared to an economy without heterogeneity, but there are selection effects that dampen the total output effect. When we eliminate those selection effects by assuming that the monetary shock is only partially observable, the total output effects are seven time larger.

In the second exercise, the monetary shock interacts with an uncertainty shock that is synchronized across all firms. We find that output responses are smaller and less persistent when average uncertainty is higher. We finish with an analytical characterization of the impulse-response function as a system of Bellman equations.

5.1 Output response to an unanticipated monetary shock

In the first exercise, we compute the impulse-response function of output to a one-time unanticipated small shock to money supply. This monetary shock is fully observed by all firms and thus we say that it is disclosed. Starting from a zero inflation steady state at $t = 0$, we shock the economy with a permanent increase in the money supply of a small size δ, such that $\log M_t = \log \bar{M} + \delta,\ t \geq 0$. Since wages are proportional to the money supply, the shock translates directly into a wage increase. In turn, the wage increase brings down all markups by δ. Given that the monetary shock is disclosed, markup estimates also fall by δ as they are updated by the full amount of the monetary shock: $\hat{\mu}_0(z) = \hat{\mu}_{-1}(z) - \delta, \quad \forall z$.

Response of aggregate price level and output Even though markup gap estimates get updated immediately, prices will only be changed when these estimates fall outside the respective inaction regions. The price index in (23) can be written in terms of the markup gaps by multiplying and dividing by the nominal wages and using the definition of markup gap:

$$P_t = W_t \left[\int_0^1 \left(\frac{p_t(z)}{W_t A_t(z)} \right)^{1-\eta} dz \right]^{\frac{1}{1-\eta}} = W_t \left[\int_0^1 \mu_t(z)^{1-\eta} dz \right]^{\frac{1}{1-\eta}} = W_t \mu^* \left[\int_0^1 \left(e^{\mu_t(z)} \right)^{1-\eta} dz \right]^{\frac{1}{1-\eta}}$$

Taking the log difference from steady state, approximating the integral, and substituting the wage deviation $\ln\left(\frac{W_t}{\overline{W}}\right) = \delta$, we obtain the price deviations from steady state denoted by $\tilde{P}_t$:

$$\tilde{P}_t \equiv \ln\left(\frac{P_t}{\overline{P}}\right) \approx \delta + \int_0^1 \mu_t(z) dz \;\approx\; \delta + \int_0^1 (\mu_t(z) - \hat{\mu}_t(z)) + \hat{\mu}_t(z)\, dz \;\approx\; \delta + \int_0^1 \hat{\mu}_t(z)\, dz \tag{29}$$

We arrive at the last equality by noticing that the forecast error $\mu_t(z) - \hat{\mu}_t(z)$ is *iid* across firms and therefore the average forecast error is equal to zero. Expression (29) states that the price level (and thus output) will deviate from its steady state value as long as some firms have not adjusted their price. To compute the output response to the monetary shock, we use the equilibrium condition that output equals the real wage. Therefore, putting together the wage and price level deviations from steady state, output deviations are given by the negative of the cross-sectional average of markup gap estimates:

$$\tilde{Y}_t \equiv \ln\left(\frac{Y_t}{\overline{Y}}\right) = \delta - \tilde{P}_t = -\int_0^1 \hat{\mu}_t(z)\, dz \tag{30}$$

We measure output effects through two statistics: the area under the impulse-response function—the total output effect— denoted by $\mathcal{M} \equiv \int_0^\infty \tilde{Y}_t\, dt$, and the half-life of the impulse response. In a frictionless world, all firms would increase their price in δ to reflect the higher marginal costs, implying that $\hat{\mu}_t(z) = 0$ for all firms and periods. The monetary shock would have no output effects. With the menu costs and the information frictions, the price level will fail to fully reflect the monetary shock and there will be real effects. During the transition to the new steady state, there are general equilibrium effects arising from changes in the average markup in the economy that affect individual policies. However in Álvarez and Lippi (2014)'s Proposition 7, it is demonstrated that in this type of framework without complementarities, such general equilibrium effects can be ignored. Following this result, we compute price responses using the steady state policies. See Section B in the Appendix for details on the computation of the steady state and the transition dynamics.

Figure VI shows the impulse-response of output to a monetary shock of size $\delta = 1\%$ for the three calibrations outlined in the previous section and Columns (1) to (3) of Table II report the total output effects and half-lives.[7]

Table II – Output Response to Monetary Shock for Different Parametrizations

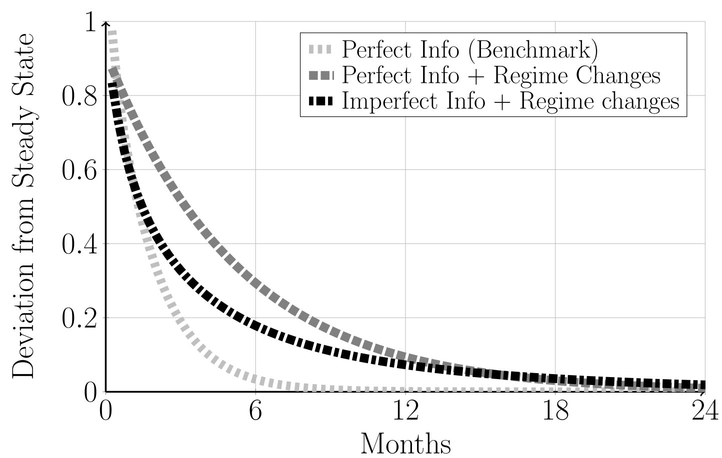
	Perfect Info		Info Frictions	
	(1)	(2)	(3)	(4)
Output Effect	Benchmark	Regime changes	Disclosed	Undisclosed
Total effect ($\mathcal{M}$)	1.00	2.85	2.14	6.98
Half-life ($t_{0.5}$)	1.00	2.67	1.00	5.33

Multiples of the benchmark case in Column (1) with perfect information and only small frequent shocks. For that case, the total output effects are $\mathcal{M} = 1.74\%$ and the half-life is 1.5 months.

[7] Note that the impulse-responses have a jump on impact. This jump arises because we solve the model in discrete time and there is a positive mass of firms at the borders of inaction; this jump does not occur in a continuous time model and the impact of a small monetary shocks is second order.

For the first calibration with only small frequent shocks (Column 1 of Table II), an increase of 1% in the money supply generates a total output effect of $\mathcal{M} = 1.74\%$, and it has a half-life of 1.5 months (these numbers are set as a benchmark). The small and short-lived output response is the result of a large selection effect as highlighted by Golosov and Lucas (2007). The firms that are more likely to adjust their price after the monetary shock are those with the largest desired price changes; their adjustments offset any potential monetary effects. The second calibration (Column 2 of Table II), which introduces the regime changes, features more than two times the total output effects and the half-life of the first model. This model generates a flat hazard rate; it is akin to a Calvo economy. By breaking the selection effect, it obtains a larger non-neutrality of monetary shocks as in Gertler and Leahy (2008) and Midrigan (2011). The third calibration with information frictions (Column 3 of Table II) doubles the output effects of the benchmark model but at the same half-life.

Figure VI – Output Impulse-Response to a Monetary Shock



Impulse-response of output for three parametrizations of the model: (1) Perfect info with only frequent shocks (benchmark case); (2) Perfect info with only infrequent shocks; (3) Imperfect info with both types of shocks.

Both the larger output effects and the shorter half-life are the result of having a large mass of firms with low uncertainty in steady state. Low uncertainty firms have small inaction regions, so the impact of the monetary shock triggers many price changes. This resembles the selection effect of the benchmark model. In fact, the adjustment frequency overshoots compared to its steady state level (see Panel D of Figure VII below) and reduces the output effect drastically during the first months. However, even with the frequency overshoot, the model with information frictions still obtains a larger output effect. The reason is that there are low uncertainty firms that did not adjust on impact, and will only incorporate the monetary shock when they receive a regime change. This delay increases the persistence.

Undisclosed monetary shock The frequency overshoot after a monetary shock is not observed in the data, as the aggregate frequency is very stable (Nakamura and Steinsson (2008), Klenow and Kryvtsov (2008)) or slightly countercyclical (Vavra (2014)). To address this issue, we consider an extension of the model where firms only observe a fraction $\alpha \in [0, 1]$ of the monetary shock, and their markup gap estimates are only partially updated[8]:

$$\hat{\mu}_0(z) = \hat{\mu}_{-1}(z) - \alpha\delta$$

[8]The CalvoPlus model with random menu costs developed in Section F of the Online Appendix also reduces the frequency overshoot and amplifies persistence.

An alternative assumption that delivers the same aggregate implications is that a random fraction of firms $\alpha \in [0, 1]$ does not observe the monetary shock. We assume that firms filter the monetary shock using the same learning technology they use to estimate their markups. Upon the impact of the monetary shock, but before the idiosyncratic shocks are realized, forecast errors $\varphi_t(z) \equiv \mu_t(z) - \hat{\mu}_t(z)$ will arise and will evolve as follows:

$$d\varphi_t(z) = -\frac{\Omega_t(z)}{\gamma}\varphi_t(z)dt + d\mathcal{G}_t(z), \quad \text{with} \quad \varphi_0(z) \sim \mathcal{N}(-(1-\alpha)\delta, \gamma\Omega_0(z)) \tag{31}$$

where we define the process $d\mathcal{G}_t(z) \equiv \sigma_f dW_t(z) + \sigma_u u_t(z)dQ_t(z) - \Omega_t(z)dZ_t(z)$ as the component of forecast errors with a cross-sectional mean of zero, i.e. $\int_0^1 \mathcal{G}_t(z)dz = 0$, since all shocks are *iid* across firms.

Note that the forecast errors of a high uncertainty firm converge faster to zero. Using (31) and its initial condition, the cross-sectional average of forecast errors is computed as

$$\mathcal{F}_t \equiv \int_0^1 \varphi_t(z)\ dz = -\int_0^1 \left[\int_0^t \frac{\Omega_s(z)}{\gamma}\varphi_s(z)ds\right] dz \tag{32}$$

and after the monetary shock it evolves as follows:

$$d\mathcal{F}_t \equiv -\int_0^1 \frac{\Omega_t(z)}{\gamma}\varphi_t(z)dz, \quad \text{with} \quad \mathcal{F}_0 = -(1-\alpha)\delta \tag{33}$$

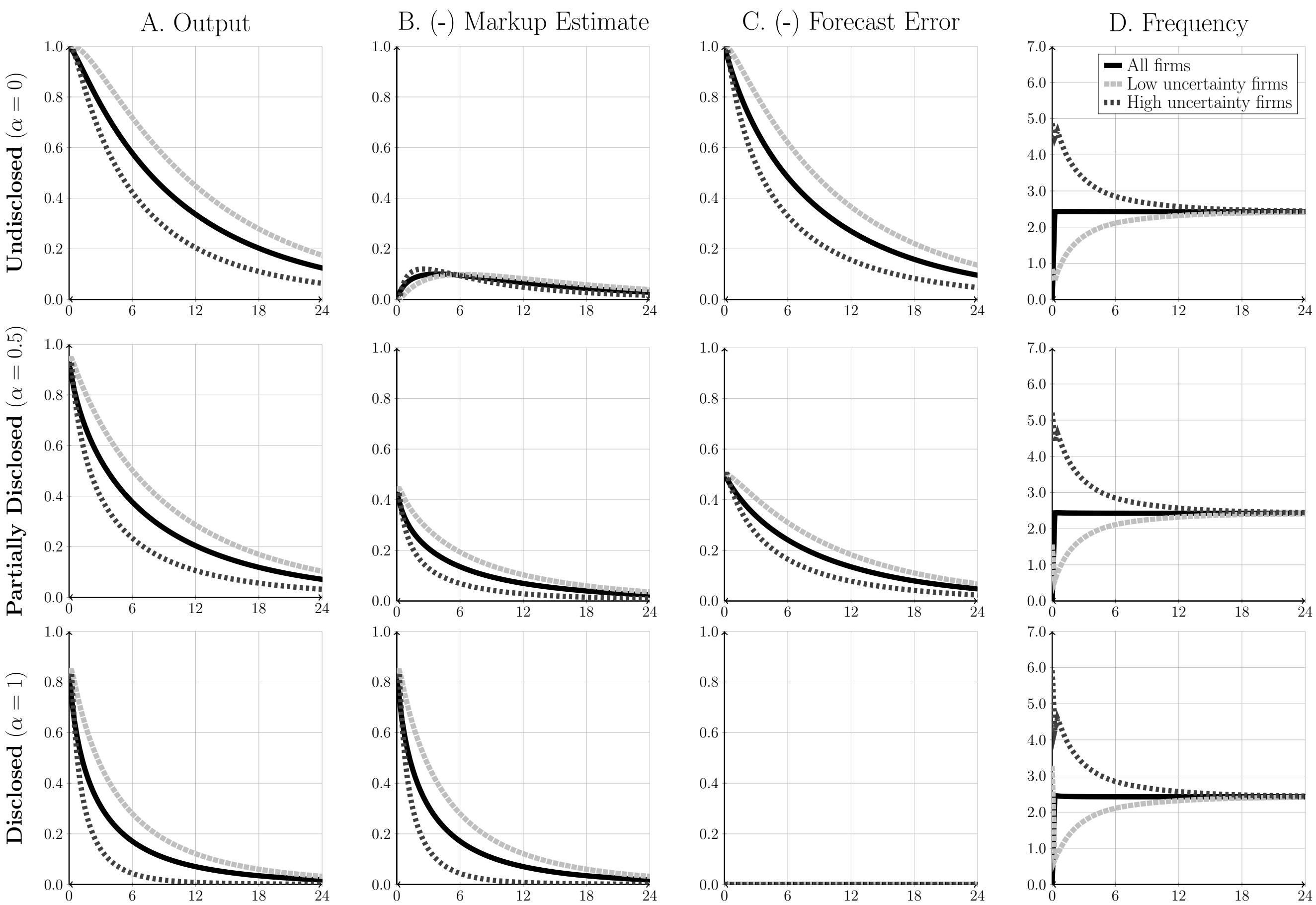
With these definitions, we can write the output deviation from steady as the negative of average markup gaps, as in (30), minus the average forecast error as follows:

$$\tilde{Y}_t = -\int_0^1 \mu_t(z)\ dz = -\int_0^1 \hat{\mu}_t(z)\ dz - \mathcal{F}_t \tag{34}$$

Contrary to the case of a disclosed monetary shock, average forecast errors will no longer be equal to zero. Furthermore, the cross-sectional distribution of uncertainty matters for the convergence rate of forecast errors towards zero. To illustrate this point, suppose there are no infrequent shocks ($\lambda = 0$) so that there is no heterogeneity in uncertainty, i.e. $\Omega_t(z) = \sigma_f,\ \forall t, z$. Then $d\mathcal{F}_t = -\frac{\sigma_f}{\gamma}\mathcal{F}_t\ dt$ and average forecast errors converge to zero at a constant exponential rate of $\frac{\sigma_f}{\gamma}$ in absolute value. In contrast, with the presence of uncertainty heterogeneity, the convergence rate of average forecast errors is given by the weighted average of individual forecast errors, where the weights are equal to individual uncertainty as in (33). Since forecast errors move inversely with uncertainty, the weighted average rate is smaller than the exponential rate. The slower convergence increases total output effects. a Figure VII plots impulse-responses of output, the average markup gap estimate, the average forecast error, and the average adjustment frequency following a monetary shock. We consider three cases for the observability of the monetary shock: undisclosed monetary shock ($\alpha = 0$), partially disclosed monetary shock ($\alpha = 0.5$), and fully disclosed monetary shock ($\alpha = 1$). In order to shed light on the role uncertainty heterogeneity, we compute cross-sectional averages *conditional on the level of uncertainty* faced by firms at the moment of the monetary shock. We display responses of all firms, firms with uncertainty below the median, and firms with uncertainty above the median. From Panel A and columns 3 and 4 of Table II, we observe that the output effect is almost tripled and the half-life quintuples when moving from disclosed to undisclosed shock. There are two forces that contribute to this amplification. First, the frequency overshoot disappears (Panel D): the adjustment frequency of low uncertainty firms that jumped with a disclosed shock now does not move on impact. Second, there is the additional persistence coming from the average forecast error (Panel C), which is the result of uncertainty heterogeneity as explained above. High uncertainty firms put a high weight on signals and incorporate the monetary shock quickly into their estimates; whereas low uncertainty firms put a low weight on signals and take a long time to incorporate the monetary shock, increasing the persistence of the average forecast error.

It is also worth noting that on impact of the monetary shock, there is no difference in the output response of high and low uncertainty firms; the difference lies in how the subsequent response is distributed between changes in the adjustment frequency and changes in the adjustment size. High uncertainty firms increase their adjustment frequency with respect to its steady state value and have on average larger markup gap estimates, while low uncertainty firms decrease their adjustment frequency and have on average smaller markup gap estimates. Since the steady state distribution of uncertainty is biased towards low values, it is the low frequency and small adjustments of the low uncertainty firms that drive the output response. For more details on the dynamics of the markup gap distribution following a monetary shock see Section G of the Online Appendix.

Figure VII – Impulse-Responses, Conditional on Firm Uncertainty and Observability of Monetary Shock



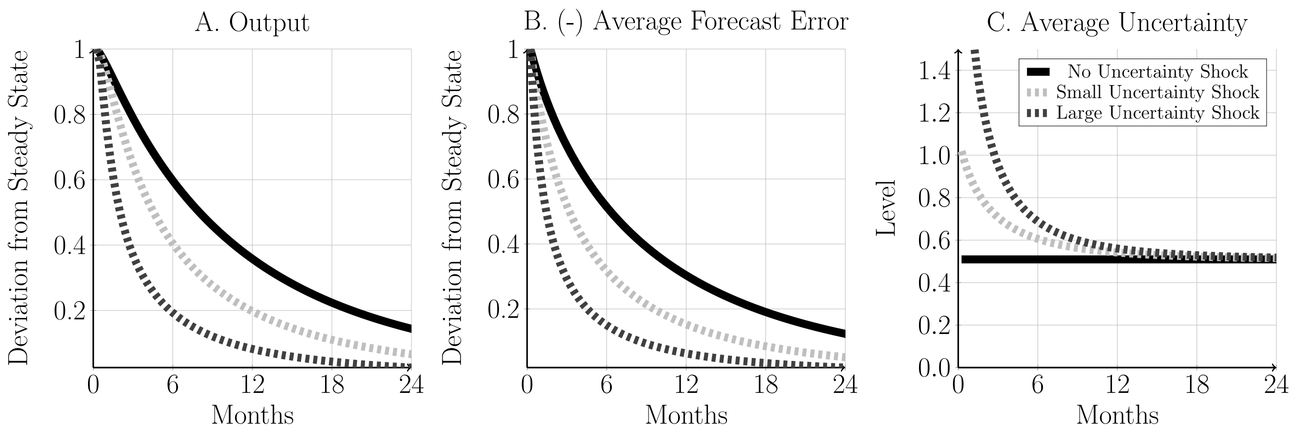
Impulse-response of output, (minus) average markup gap estimate, (minus) average forecast error and average adjustment frequency after a monetary shock. The first three variables are measured as deviations from steady state, while adjustment frequency is plotted in levels. Column A is equal to the sum of Columns B and C. Observability of monetary shock: first row = undisclosed, second row = partially disclosed, third row = disclosed. Responses are conditional on the initial level of uncertainty: solid line = total mass of firms, light dashed line = uncertainty below the median, dark dotted dashed line = uncertainty above the median.

5.2 Aggregate uncertainty and nominal shocks

The second exercise explores the output response to a monetary shock when it occurs at the same time as an aggregate uncertainty shock. The motivation for this exercise is to provide an explanation for the empirical finding that monetary policy is less effective when economic uncertainty is higher. While modeling aggregate uncertainty shocks is outside the scope of this paper, we analyze the interaction of an undisclosed monetary shock with a one-time exogenous and unanticipated uncertainty shock synchronized across firms. The uncertainty shock increases every firm's uncertainty by $\kappa\bar{\Omega}$, where $\bar{\Omega}$ is average steady state uncertainty and $\kappa \in \{0, 1, 4\}$. An example of this type of shock is a monetary expansionary during a recession or any period of elevated economic uncertainty.

Figure VIII shows the output impulse-response, the average forecast error, and average uncertainty for each experiment and Table III reports the statistics.

Figure VIII – Impulse-Responses to Monetary Shock and Synchronized Uncertainty Shock



Impulse-response of output, (minus) average forecast error, and average uncertainty after a monetary shock. The first two variables are measured as deviations from steady state, while uncertainty is plotted in levels. Higher average uncertainty reduces the output effects from the monetary shock.

Table III – Output Effects of Monetary and Synchronized Uncertainty Shock

Output Effect	No Ω shock $\kappa = 0$	Small Ω shock $\kappa = 1$	Large Ω shock $\kappa = 4$
Total effect ($\mathcal{M}$)	6.98	4.51	2.42
Half-life ($t_{0.5}$)	5.33	3.06	1.45

As multiples of benchmark case, reported in Column (1) in Table II.

Panel A shows that a monetary shock paired with a small uncertainty shock reduces the output response and half-life in 40%; and if it is paired with a high uncertainty shock, the output effects are significantly reduced. The positive relationship between adjustment frequency and uncertainty is also present here: higher firm uncertainty makes the aggregate price level more flexible and decreases output effects. This effect is also present in Vavra (2014), where aggregate volatility shocks are explicitly modeled.

There is an additional effect that is particular to our model and has to do with forecast error dynamics. In more uncertain times, firms place a higher weight on new information, forecast errors disappear faster, and the monetary shock is quickly incorporated into prices; this reduces the persistence of the average forecast error, and in turn, the persistence of the output response. This can be seen in Panel B, which shows that the average forecast error $\mathcal{F}_t$ converges faster to zero when uncertainty is higher.

Finally, in Panel C we observe that the uncertainty shocks are short-lived, as average uncertainty converges back to its steady state level after a few months. The magnitude and persistence of uncertainty shocks are comparable to the ones documented in Bloom (2009) and Jurado, Ludvigson and Ng (2015).

The relationship between aggregate uncertainty and forecast errors is novel and there is empirical evidence that supports it. Coibion and Gorodnichenko (2015) compares the dynamics of forecast errors during periods of high economic volatility (such as the 70's and 80's) with periods of low economic volatility (such as the late 90's). It concludes that information rigidities are higher during periods of low uncertainty than higher uncertainty, just as our model predicts.

5.3 Characterization of the impulse-response function

To summarize all the previous results, we characterize analytically the total output effect after aggregate monetary and uncertainty shocks. Following the strategy in Álvarez, Le Bihan and Lippi (2014), the next proposition expresses the output effect as a system of Bellman equations.

Proposition 9 (**Output Effects from Monetary and Uncertainty Shocks**). *Assume the economy is in steady state and it is hit with one-time unanticipated monetary shock of size δ, and firms only observe a fraction $\alpha \in [0,1]$ of it. Simultaneously, idiosyncratic firm uncertainty increases by $\kappa\bar{\Omega}$. Before the monetary and uncertainty aggregate shocks, firms' states are denoted by $(\hat{\mu}_{-1}, \Omega_{-1})$ distributed according to F.*

1. *Immediately after aggregate shocks arrive, but before idiosyncratic shocks do, markup estimates and uncertainty jump to $\hat{\mu}_0 = \hat{\mu}_{-1} - \alpha\delta$ and $\Omega_0 = \Omega_{-1} + \kappa\bar{\Omega}$. Before idiosyncratic shocks hit, forecast errors are random, and conditional on uncertainty, they are Normally distributed: $\varphi_0 \sim \mathcal{N}(-(1-\alpha)\delta, \gamma\Omega_0)$.*

2. *Let w be the future stream of pricing mistakes for a firm with state $(\hat{\mu}, \Omega, \varphi)$; it is computed recursively as*

$$w(\hat{\mu}, \Omega, \varphi) \quad = \quad \mathbb{E}\left[\int_0^\tau (\hat{\mu}_t + \varphi_t)\, dt + w(0, \Omega_\tau, \varphi_\tau) \Big| (\hat{\mu}_0, \Omega_0, \varphi_0) = (\hat{\mu}, \Omega, \varphi)\right] \tag{35}$$

subject to the following stochastic process:

$$d\hat{\mu}_t = \Omega_t \frac{\varphi_t}{\gamma} dt + \Omega_t dZ_t; \quad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t; \quad d\varphi_t = -\Omega_t \frac{\varphi_t}{\gamma} dt + \sigma_f dW_t + \sigma_u u_t dQ_t - \Omega_t dZ_t$$

3. *The total output response averages across all firms streams of pricing mistakes, taking into account the steady state distribution and the distribution of forecast errors:*

$$\mathcal{M}(\delta, \alpha, \kappa\bar{\Omega}) = -\int_{\hat{\mu},\Omega} \left[\int_{\varphi_0} w(\hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, x)\phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma(\Omega_{-1} + \kappa\bar{\Omega})}\right) d\varphi_0\right] dF(\hat{\mu}_{-1}, \Omega_{-1}) \tag{36}$$

4. *If $\alpha = 1$ (fully disclosed), then*

$$\mathcal{M}(\delta, 1, \kappa\bar{\Omega}) = -\int_{\hat{\mu},\Omega} \mathbb{E}\left[\int_0^\tau \hat{\mu}_t\, dt \Big| (\hat{\mu}_0, \Omega_0) = (\hat{\mu}_{-1} - \delta, \Omega_{-1} + \kappa\bar{\Omega})\right] dF(\hat{\mu}_{-1}, \Omega_{-1}) \tag{37}$$

subject to: $\quad d\hat{\mu}_t = \Omega_t d\hat{Z}_t; \quad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t.$

Pricing mistakes computed in (35) arise either from markup gap estimates that fall inside the inaction region—the firm is aware of these mistakes, which are optimal—or from forecast errors that delay the update of markup gap estimates. Equation (36) averages the pricing mistakes across all firms. This result makes evident that pricing mistakes do not disappear after a firm's first price change, as it happens when the monetary shock is disclosed, but they persist as the monetary shock is only partially incorporated with each subsequent adjustment. Notice that uncertainty heterogeneity affects the output response through the stochastic process of both markup gaps and forecast errors: it increases the dispersion of expected times across ex-ante identical firms, and it decreases the speed of convergence of forecast errors.

In the case of a disclosed shock ($\alpha = 1$), average forecast errors are equal to zero because innovations are *iid* across firms and heterogeneity only plays through dispersion in expected times. The *first* price change of each firm fully incorporates the monetary shock, and for this reason, equation (37) does not have a recursive nature.

6 Uncertainty and pass-through

In the previous exercises, we established connections between the dynamics of aggregate uncertainty, aggregate forecast errors, and the propagation of monetary shocks. In this section, we study these connections at the individual level using the concept of pass-through, which measures the responsiveness of individual prices to the monetary shock. We establish two results. First, we find that information frictions, disciplined via micro-price data, reduce pass-through. Second, we find a positive relationship between firm uncertainty and pass-through.

We follow the methodology used to estimate the pass-though of nominal shocks into prices as in Gopinath, Itskhoki and Rigobón (2010). We consider a random walk process for the log deviations of money supply from its steady state, $\ln M_{t+1} = \ln M_t + \sigma_M \varepsilon_{t+1}$, with $\varepsilon_{t+1} \sim \mathcal{N}(0,1)$, with volatility $\sigma_M = 0.007$ at weekly frequency. We then generate a panel of prices for $N = 10,000$ firms denoted with i and for $T = 100,000$ periods denoted with t. For each firm, we record the size of the price change Δp_t^i and the cumulative nominal shock $\Delta^c M_t^i$ measured as the money supply deviations from steady state between her price changes: $\Delta^c M_t^i \equiv \ln M_t - \ln M^i_{t-n_t^i}$, where n_t^i denotes the number of periods since her last price change.

We regress the size of price changes into the cumulative monetary shock $\Delta_c M_t^i$, firm uncertainty Ω_t^i, and an interaction term between firm uncertainty and the monetary shock ($\Delta_c M_t^i \times \Omega_t^i$).

$$\Delta p_t^i = \beta_M \; \Delta_c M_t^i + \beta_\Omega \; \Omega_t^i + \beta_{M,\Omega} \; (\Delta_c M_t^i \times \Omega_t^i) + \epsilon_t^i \tag{38}$$

Results from different specifications of this regression are reported in Table IV. Columns (1) to (3) report results for three models with perfect information: Calvo; menu cost with only frequent shocks; and menu cost with only infrequent shocks. In these cases, the coefficient on the cumulative monetary shock β_M measures the average pass-through of the nominal shock into the price. Unsurprisingly, we find that pass-though is complete (β_M is very close to unity): conditional on a price change, the firms fully incorporate the money shock into their prices.

The last four columns report results for our model with heterogenous firm uncertainty, with and without disclosed money shocks. When we include the interaction term, the average pass-through is measured by $\beta_M + \beta_{M,\Omega} \; \bar{\Omega}$, where average uncertainty level is $\bar{\Omega} = 0.056$. In the model with a disclosed monetary shock in Columns (4a) and (4b), average pass-through is equal to 1.02 or 1.05 if we include the interaction; thus pass-through is complete. In the model with undisclosed monetary shock in Columns (5a) and (5b), average pass-through is equal to 0.23 or 0.20 if we include the interaction term; thus it is five times smaller. The information friction delays the updating of the permanent component of marginal costs. This is a success of the model as it brings the pass-through coefficient closer to the small numbers found in the data.

Uncertainty on its own is not statistically significant, but its interaction with the cumulative money shock yields

Table IV – Firm Uncertainty and Nominal Pass-Through

Regressor	Coefficient	(1)	(2)	(3)	(4a)	(4b)	(5a)	(5b)
Monetary shock	β_M	1.00	1.03	1.02	1.02	1.12	0.23	0.18
		(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.02)
Uncertainty	β_Ω					0.02		-0.05
						(0.04)		(0.04)
Interaction	$\beta_{M,\Omega}\bar{\Omega}$					-0.07		0.02
						(0.01)		(0.01)
R^2		0.14	0.14	0.11	0.005	0.005	0.0003	0.0003

Robust standard errors in parenthesis. The interaction term is evaluated at average uncertainty $\bar{\Omega} = 0.056$. Models: (1) Calvo; (2) Perfect info with only frequent shocks; (3) Perfect info with only infrequent shocks; (4) Imperfect info and disclosed monetary shock; and (5) Imperfect info and undisclosed monetary shock.

very interesting results. When monetary shocks are observable in an environment of uncertain idiosyncratic productivity (Column 4b), the coefficient $\beta_{M,\Omega}$ is negative: when firm uncertainty is high, idiosyncratic productivity shocks become relatively more important than monetary shocks for pricing decisions; this reduces selection effects and average pass-through. In contrast, when the monetary shocks are unobservable (Column 5b), the coefficient $\beta_{M,\Omega}$ is positive. In this case, highly uncertain firms assign a larger Bayesian weight to observations that contain the monetary shock, and incorporate a larger fraction of the shock into their prices. Given the positive relationship between uncertainty and standard deviation of price changes in Proposition 6, our results imply a positive relationship between the standard deviation of price changes and pass-through, as documented in Berger and Vavra (2015) in the context of import price-setting.

The low pass-through of nominal shocks into individual prices is often attributed to strategic complementarities across firms. Our results show that information frictions about the nominal shock is an alternative way to decrease pass-through. Complementarities achieve it by decreasing the elasticity of the size of price changes with respect to nominal marginal costs, whereas information frictions achieve it because firms take time to realize that costs have changed. However, the two mechanisms make opposite predictions regarding the relationship between idiosyncratic uncertainty and pass-through. For instance, Berger and Vavra (2015) shows that in model with strategic complementarities that arise from Kimball demand, larger volatility of idiosyncratic shocks reduces pass-through; while in our model with undisclosed nominal shocks, larger uncertainty about idiosyncratic shocks increases pass-through. This appears as an interesting implication for empirical research.

7 Conclusions

In this paper we develop a framework to analyze pricing policies in environments with idiosyncratic uncertainty, as well as the role of heterogeneous uncertainty in amplifying the effects of nominal shocks. We show that the combination of menu costs and uncertainty cycles can generate persistent output responses while also explaining micro evidence on decreasing hazard rates and age dependent price statistics. Furthermore, this combination can explain recent evidence regarding the effectiveness of monetary policy during highly uncertain times, as well as the way in which uncertainty shapes forecast error dynamics.

Our model combines an inaction problem arising from a non-convex adjustment cost together with a signal extraction problem with jumps, where agents face undistinguishable transitory and permanent shocks. As far as we know, our paper is the first to solve this problem type analytically and deliver predictions for the joint dynamics of uncertainty, actions, and forecast errors. Although the focus here is on pricing decisions, the model is easy to generalize to other setups where fixed adjustment costs, large infrequent shocks, and information frictions are likely to coexist. For example, it could be applied to analyze portfolio allocation problems subject to adjustment fees and a stochastic trend in the dividend payment, or to study the problem of a worker that decides whether to shift occupations subject to a mobility cost and uncertain productivity growth. Particularly, we foresee applications in setups that generate strong age dependent statistics, such as labor markets. Moreover, the tractability of our filtering framework with regime changes could facilitate the study of disaster risk impact on asset prices, where agents know a disaster has happened but do not know the true magnitude of its effects.

Going forward, it would be interesting to explore our model's implications for state dependency in filtering and pricing decisions. Do sectors with more heterogeneous uncertainty—measured with price statistics—feature lower forecast errors persistence? Does monetary policy have smaller effects in more uncertain countries? Should monetary policy rules internalize firm uncertainty and how? New surveys about firm expectations, as in Coibion, Gorodnichenko and Kumar (2015), together with insights from our model, could help address these questions.

References

Aastveit, K. A., Natvik, G. J. and Sola, S. (2013). Economic uncertainty and the effectiveness of monetary policy. *Norges Bank Working Paper 17*.

Álvarez, F., Le Bihan, H. and Lippi, F. (2014). *Small and Large Price Changes and the Propagation of Monetary Shocks*. Working Paper 20155, National Bureau of Economic Research.

— and Lippi, F. (2014). Price setting with menu cost for multi-product firms. *Econometrica*, **82** (1), 89–135.

Álvarez, F. and Lippi, F. (2015). *Price plans and the real effects of monetary policy*. Tech. rep.

Álvarez, F., Lippi, F. and Paciello, L. (2011). Optimal price setting with observation and menu costs. *The Quarterly Journal of Economics*, **126** (4), 1909–1960.

Álvarez, F., Lippi, F. and Paciello, L. (2016). Monetary shocks in models with inattentive producers. *The Review of Economic Studies*, **83** (2), 421–459.

Álvarez, L. J., Burriel, P. and Hernando, I. (2005). Do decreasing hazard functions for price changes make any sense? *ECB Working Paper*.

Argente, D. and Yeh, C. (2015). *A Menu Cost Model with Price Experimentation*. Tech. rep., Working paper.

Bachmann, R., Born, B., Elstner, S. and Grimme, C. (2013). *Time-Varying Business Volatility, Price Setting, and the Real Effects of Monetary Policy*. Working Paper 19180, National Bureau of Economic Research.

— and Moscarini, G. (2011). *Business Cycles and Endogenous Uncertainty*. 2011 Meeting Papers 36, Society for Economic Dynamics.

Baley, I., Kochen, F. and Sámano, D. (2016). Price duration and pass-through. *Mimeo, Banco de México*.

Barro, R. J. (1972). A theory of monopolistic price adjustment. *The Review of Economic Studies*, **39** (1), 17–26.

Berger, D. and Vavra, J. S. (2015). *Volatility and Pass-through*. Working Paper 19651, National Bureau of Economic Research.

Blanco, A. J. (2016a). *Costs and Benefits of Inflation in Models with Price Rigidities*. Working paper, University of Michigan.

— (2016b). *Optimal Inflation Target in an Economy with Menu Costs and an Occasionally Binding Zero Lower Bound*. Working paper, University of Michigan.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, **77** (3), 623–685.

Borovičková, K. (2013). *What Drives Labor Market Flows?* Working paper, New York University, Mimeo.

Caggiano, G., Castelnuovo, E. and Nodari, G. (2014). *Uncertainty and Monetary Policy in Good and Bad Times*. Tech. rep., Dipartimento di Scienze Economiche" Marco Fanno".

Campbell, J. R. and Eden, B. (2014). Rigid prices: Evidence from us scanner data. *International Economic Review*, **55** (2), 423–442.

Caplin, A. and Leahy, J. (1997). Aggregation and optimization with state-dependent pricing. *Econometrica: Journal of the Econometric Society*, pp. 601–625.

Carvalho, C. (2006). Heterogeneity in price stickiness and the real effects of monetary shocks. *The BE Journal of Macroeconomics*, **6** (3), 1–58.

— and Schwartzman, F. (2015). Selection and monetary non-neutrality in time-dependent pricing models. *Journal of Monetary Economics*, **76**, 141 – 156.

Chu, M. B., Huynh, K. P., Jacho-Chavez, D. and Kryvtsov, O. (2016). *On the Evolution of the United Kingdom Price Distributions*. Tech. rep.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, **105** (8), 2644–78.

—, — and Kumar, S. (2015). *How do firms form their expectations? new survey evidence*. Tech. rep., National Bureau of Economic Research.

Cortés, J., Murillo, J. A. and Ramos-Francia, M. (2012). Evidencia de los micro datos del inpc respecto al proceso de formación de precios. *manuscript, Banco de México*.

Davis, M. H. (1977). Linear estimation and stochastic control. *Press, London*.

Dhyne, E., Alvarez, L. J., Le Bihan, H., Veronese, G., Dias, D., Hoffmann, J., Jonker, N., Lunnemann, P., Rumler, F. and Vilmunen, J. (2006). Price changes in the euro area and the united states: Some facts from individual consumer price data. *The Journal of Economic Perspectives*, **20** (2), 171–192.

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

Eden, B. and Jaremski, M. S. (2009). *Rigidity, Dispersion and Discreteness in Chain Prices*. Working paper, Vanderbilt University.

Gertler, M. and Leahy, J. (2008). A phillips curve with an ss foundation. *Journal of Political Economy*, **116** (3), 533–572.

Golosov, M. and Lucas, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, pp. 171–199.

Gopinath, G., Itskhoki, O. and Rigobón, R. (2010). Currency choice and exchange rate pass-through. *The American economic review*, **100** (1), 304–336.

Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, **57** (2), 357–84.

Hellwig, C. and Venkateswaran, V. (2009). Setting the right prices for the wrong reasons. *Journal of Monetary Economics*, **56**, S57–S77.

Jovanovic, B. (1979). Job Matching and the Theory of Turnover. *Journal of Political Economy*, **87** (5), 972–90.

Jurado, K., Ludvigson, S. C. and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, **105** (3), 1177–1216.

Karadi, P. and Reiff, A. (2014). Menu costs, aggregate fluctuations, and large shocks. *CEPR Discussion Paper No. DP10138*.

Kehoe, P. and Midrigan, V. (2015). Prices are sticky after all. *Journal of Monetary Economics*, **75**, 35–53.

Keller, G. and Rady, S. (1999). Optimal experimentation in a changing environment. *The Review of Economic Studies*, **66** (3), 475–507.

Kiefer, N. M. (1988). Economic duration data and hazard functions. *Journal of economic literature*, **26** (2), 646–679.

Kim, C.-J. (1994). Dynamic linear models with markov-switching. *Journal of Econometrics*, **60** (1-2), 1–22.

Klenow, P. J. and Kryvtsov, O. (2008). State-dependent or time-dependent pricing: Does it matter for recent u.s. inflation? *Quarterly Journal of Economics*, **123** (3), 863–904.

Kolkiewicz, A. W. (2002). Pricing and hedging more general double-barrier options. *Journal of Computational Finance*, **5** (3), 1–26.

Kryvtsov, O. and Vincent, N. (2016). *The Cyclicality of Sales for Aggregate Price Flexibility*. Tech. rep.

Kushner, H. J. and Dupuis, P. G. (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 2nd edn.

Levy, D., Bergen, M., Dutta, S. and Venable, R. (1997). The magnitude of menu costs: Direct evidence from large u. s. supermarket chains. *The Quarterly Journal of Economics*, **112** (3), 791–824.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory*, **4** (2), 103–124.

Maćkowiak, B. and Wiederholt, M. (2009). Optimal Sticky Prices under Rational Inattention. *American Economic Review*, **99** (3), 769–803.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, **117** (4), 1295–1328.

Matějka, F. (2015). Rationally inattentive seller: Sales and discrete pricing. *The Review of Economic Studies*.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, **3** (1), 125–144.

Midrigan, V. (2011). Menu cost, multiproduct firms, and aggregate fluctuations. *Econometrica*, **79** (4), 1139–1180.

Mumtaz, H. and Surico, P. (2015). The transmission mechanism in good and bad times. *International Economic Review*, **56** (4), 1237–1260.

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

— and — (2010). Monetary non-neutrality in a multisector menu cost model. *The Quarterly journal of economics*, **125** (3), 961–1013.

Øksendal, B. (2007). *Stochastic Differential Equations*. Springer, 6th edn.

Øksendal, B. K. and Sulem, A. (2010). *Applied stochastic control of jump diffusions*, vol. 498. Springer.

Pellegrino, G. (2014). Uncertainty and monetary policy in the us: A journey into non-linear territory. *University of Verona, mimeo*.

Senga, T. (2016). *A New Look at Uncertainty Shocks: Imperfect Information and Misallocation*. Working paper, Queen Mary University of London.

Stokey, N. (2009). *The Economics of Inaction*. Princeton University Press.

Tenreyro, S. and Thwaites, G. (2015). Pushing on a string: Us monetary policy is less powerful in recessions. *CEPR Discussion Paper No. DP10786*.

Vavra, J. (2010). The empirical price duration distribution and monetary non-neutrality. *Working paper series, SSRN eLibrary*.

— (2014). Inflation dynamics and time-varying volatility: New evidence and an ss interpretation. *The Quarterly Journal of Economics*, **129** (1), 215–258.

Willems, T. (2013). Actively learning by pricing: A model of an experimenting seller. *Available at SSRN 2381430*.

Woodford, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics*, **56**, S100–S124.

Zbaracki, M. J., Ritson, M., Levy, D., Dutta, S. and Bergen, M. (2004). Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets. *The Review of Economics and Statistics*, **86** (2), 514–533.

A Appendix: Proofs

Notation Throughout the proofs, we denote partial derivatives with $f_{\hat{\mu}^i\Omega^j} \equiv \frac{\partial^{i+j} f}{\partial \hat{\mu}^i \partial \Omega^j}$.

Preliminaries: Infinitesimal generator and its adjoint operator

($\mathcal{A}$) **Infinitesimal generator.** The infinitesimal generator of $(\hat{\mu}, \Omega)$ denoted by $\mathcal{A}$, applied to a continuous bounded function ϕ is given by

$$\mathcal{A}\phi(X(t)) \equiv \lim_{dt \downarrow 0} \frac{\mathbb{E}[\phi(X(t+dt)) - \phi(X(t))]}{dt}$$

For our problem, the generator is given by:

$$\mathcal{A}\phi(\hat{\mu}_t, \Omega_t) = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} \phi_\Omega(\hat{\mu}_t, \Omega_t) + \frac{\Omega_t^2}{2} \phi_{\hat{\mu}^2}(\hat{\mu}_t, \Omega_t) + \lambda \left[\phi\left(\hat{\mu}_t, \Omega_t + \frac{\sigma_u^2}{\gamma}\right) - \phi(\hat{\mu}_t, \Omega_t) \right] \tag{A.1}$$

Note: A key property of our generator $\mathcal{A}$ is the lack of interaction terms between uncertainty and markup gap estimates. This property is implied by the passive learning process in which the firm cannot change the quality of the information flow by changing her markup.

Proof. First we need to get a formula for a jump-diffusion process analogous to Itø's formula that computes changes in $\phi(X(t))$. We follow the general description in Theorem 1.16 of Øksendal and Sulem (2010). Let $B(t)$ be an m-dimensional Brownian motion and $\{N(dt)\}$ are l independent Poisson random measures each with intensity λ_j. Then consider a multidimensional Itō-Lévy process $X(t)$, where each component is given by

$$dX_i(t) = \alpha_i(t)dt + \sum_{j=1}^{m} \sigma_{ij}(t) dB_j(t) + \sum_{j=1}^{l} \int_{\mathbb{R}} \gamma_{ij}(t) N_j(dt)$$

Let $X^c(t)$ be the continuous part of $X(t)$ (obtained by removing the jumps). Changes in $\phi(X(t))$ arise from increments in $X^c(t)$ plus the jumps coming from $N(dt)$:

$$\begin{aligned} \phi(X(t+dt)) - \phi(X(t)) &= \frac{\partial \phi}{\partial t}(t, X(t))dt + \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}(t, X(t))[\alpha_i(t)dt + \sigma_i(t)dB_t] + \frac{1}{2}\sum_{i,j=1}^{n} (\sigma\sigma')_{ij}(t) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t, X(t))dt \\ &+ \sum_{k=1}^{l} \int_{\mathbb{R}} \left\{ [\phi(t, X(t^-) + \gamma^k(t)] - \phi(t, X(t^-)) \right\} N_k(dt) \end{aligned}$$

where γ^k is column k of the $n \times l$ matrix γ and σ_i is row i of σ. To obtain the generator $\mathcal{A}$, take expectations of the previous formula (note that $\mathbb{E}[dB_t] = 0$ and $\mathbb{E}[N_j(dt)] = \lambda_j dt$), divide by dt and take the limit as $dt \to 0$, yields:

$$\begin{aligned} \mathcal{A}\phi(X(t)) \equiv \lim_{dt \downarrow 0} \frac{\mathbb{E}[\phi(X(t+dt)) - \phi(X(t))]}{dt} &= \frac{\partial \phi}{\partial t}(t, X(t)) + \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}(t, X(t))\alpha_i(t) + \frac{1}{2}\sum_{i,j=1}^{n} (\sigma\sigma')_{ij}(t) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t, X(t)) \\ &+ \sum_{k=1}^{l} \lambda_j \left\{ [\phi(t, X(t^-) + \gamma^k(t)] - \phi(t, X(t^-)) \right\} \end{aligned}$$

To apply this formula in our context, use the following relationships to obtain formula (A.1):

$$X(t) = [\hat{\mu}_t,\ \Omega_t]',\ B(t) = [d\hat{Z}_t,\ 0]',\ N(t) = [0\ q(t)]',\ \alpha_1(t) = 0,\ \alpha_2(t) = \frac{\sigma_f^2 - \Omega^2}{\gamma},\ \sigma_{11}(t) = \Omega_t,\ \gamma_{11}(t) = \frac{\sigma_u^2}{\gamma}$$

and all other entries equal to zero. Also, since we will work in a stationary environment, we set $\frac{\partial \phi}{\partial t}(t, X(t)) = 0$. □

($\mathcal{A}^*$) **Adjoint operator.** The adjoint of $\mathcal{A}$, denoted by $\mathcal{A}^*$, is such that $< \mathcal{A}\phi, f > = < \phi, \mathcal{A}^* f >$, where $<,>$ denotes the $\mathcal{L}^2$-inner product. It is given by

$$\mathcal{A}^* f(\hat{\mu},\Omega) = -\frac{\sigma_f^2-\Omega^2}{\gamma} f_\Omega(\hat{\mu},\Omega) + \frac{2\Omega}{\gamma} f(\hat{\mu},\Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu},\Omega) + \lambda \left[f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu},\Omega)\right] \tag{A.2}$$

Proof. To obtain the adjoint operator, let us apply the definition.

$$< \mathcal{A}\phi, f > = \int_{\sigma_f}^{\infty} \int_{|\hat{\mu}| \le \bar{\mu}(\Omega)} \left\{ \frac{\sigma_f^2-\Omega^2}{\gamma} \phi_\Omega(\hat{\mu},\Omega) + \frac{\Omega^2}{2} \phi_{\hat{\mu}^2}(\hat{\mu},\Omega) + \lambda \left[\phi\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - \phi(\hat{\mu},\Omega)\right]\right\} f(\hat{\mu},\Omega) d\hat{\mu} d\Omega$$

Let us simplify each integral and isolate $\phi(\hat{\mu},\Omega)$ from other terms. We highlight it in bold to make it easier to track.

(i) The first integral is computed by integration by parts with respect to Ω. We also assume that $\lim_{x\to\infty} \phi(\hat{\mu}, x) = 0$.

$$\begin{aligned}
\int\int \phi_\Omega(\hat{\mu},\Omega) \frac{\sigma_f^2-\Omega^2}{\gamma} f(\hat{\mu},\Omega) d\hat{\mu} d\Omega &= \underbrace{\int \phi(\hat{\mu},x) \frac{\sigma_f^2-x^2}{\gamma} f(\hat{\mu},x) \Big|_{\sigma_f}^{\infty} d\hat{\mu}}_{=0} - \int\int \frac{\partial}{\partial\Omega}\left(\frac{\sigma_f^2-\Omega^2}{\gamma} f(\hat{\mu},\Omega)\right) \phi(\hat{\mu},\Omega) d\hat{\mu} d\Omega \\
&= \int\int \left(-\frac{\sigma_f^2-\Omega^2}{\gamma} f_\Omega(\hat{\mu},\Omega) + \frac{2\Omega}{\gamma} f(\hat{\mu},\Omega)\right) \boldsymbol{\phi(\hat{\mu},\Omega)} d\hat{\mu} d\Omega
\end{aligned}$$

(ii) The second integral is computed integrating by parts twice with respect to $\hat{\mu}$:

$$\begin{aligned}
\int\int \frac{\Omega^2}{2} \phi_{\hat{\mu}^2}(\hat{\mu},\Omega) f(\hat{\mu},\Omega) d\hat{\mu} d\Omega &= \int \frac{\Omega^2}{2} \left[f(x,\Omega)\phi_{\hat{\mu}}(x,\Omega) - f_{\hat{\mu}}(x,\Omega)\phi(x,\Omega) \Big|_{-\bar{\mu}(\Omega)}^{\bar{\mu}(\Omega)} + \int f_{\hat{\mu}^2}(\hat{\mu},\Omega)\phi(\hat{\mu},\Omega) d\hat{\mu}\right] d\Omega \\
&= \int\int \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu},\Omega) \boldsymbol{\phi(\hat{\mu},\Omega)} d\hat{\mu} d\Omega
\end{aligned}$$

where the first term is equal to zero since $f(\bar{\mu}(\Omega),\Omega) = f(-\bar{\mu}(\Omega),\Omega) = 0$ and $\phi(\bar{\mu}(\Omega),\Omega) = \phi(-\bar{\mu}(\Omega),\Omega) = 0$.

(iii) For the third integral, we split the Ω domain in two disjoint sets and use a change of variable to rewrite it as:

$$\begin{aligned}
\int\int \lambda \left[\phi\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - \phi(\hat{\mu},\Omega)\right] f(\hat{\mu},\Omega) d\hat{\mu} d\Omega &= \int_{\sigma_f+\frac{\sigma_u^2}{\gamma}}^{\infty} \int_{|\hat{\mu}|\le\bar{\mu}(\Omega)} \lambda \left[f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu},\Omega)\right] \phi(\hat{\mu},\Omega) d\hat{\mu} d\Omega \\
&- \int_{\sigma_f}^{\sigma_f+\frac{\sigma_u^2}{\gamma}} \int_{|\hat{\mu}|\le\bar{\mu}(\Omega)} f(\hat{\mu},\Omega)\phi(\hat{\mu},\Omega) d\hat{\mu} d\Omega \\
&= \int\int \lambda \left[f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu},\Omega)\right] \boldsymbol{\phi(\hat{\mu},\Omega)} d\hat{\mu} d\Omega
\end{aligned}$$

For the second equality, notice that f's second argument only takes positive values. We define f to be equal to zero outside its domain, and therefore $f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right)\phi(\hat{\mu},\Omega) = 0$ for all $\Omega \in [\sigma_f, \sigma_f + \frac{\sigma_u^2}{\gamma}]$. Therefore, we can add the missing terms and integrate over the complete domain.

Putting all the integrals together we recover the adjoint operator A^*:

$$\int\int \underbrace{\left\{ -\frac{\sigma_f^2-\Omega^2}{\gamma} f_\Omega(\hat{\mu},\Omega) + \frac{2\Omega}{\gamma} f(\hat{\mu},\Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu},\Omega) + \lambda \left[f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu},\Omega)\right]\right\}}_{A^*} \boldsymbol{\phi(\hat{\mu},\Omega)} d\hat{\mu} d\Omega = < \phi, \mathcal{A}^* f >$$

□

Proposition 1 is proved in a more general setup than in the text, adding a non-zero drift for the state.

Proposition 1 (**Filtering Equations, Including Drift**). *Let the following processes define the state and the signal*

$$\begin{aligned}
(state) \quad d\mu_t &= F\mu_t dt + \sigma_f dW_t + \sigma_u u_t dQ_t \\
(observation) \quad ds_t &= G\mu_t dt + \gamma dZ_t \\
(initial\ conditions\ for\ state) \quad \mu_0 &\sim \mathcal{N}(a, b) \\
(initial\ conditions\ for\ observations) \quad s_0 &= 0 \\
where \quad W_t, Z_t \sim Wiener\ Process, \quad Q_t &\sim Poisson(\lambda), \quad u_t \sim \mathcal{N}(0,1)
\end{aligned} \tag{A.3}$$

Let the information set (with continuous sampling) be $\mathcal{I}_t = \sigma\{s_h, Q_h : h \in [0,t]\}$. *Then the posterior distribution of the state is Normal, i.e.* $\mu_t|\mathcal{I}_t \sim \mathcal{N}(\hat{\mu}_t, \Sigma_t)$, *where the posterior mean* $\hat{\mu}_t \equiv \mathbb{E}[\mu_t|\mathcal{I}_t]$ *and posterior variance* $\Sigma_t \equiv \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|\mathcal{I}_t]$ *satisfy the following stochastic processes:*

$$\begin{aligned}
d\hat{\mu}_t &= \left(F - \tfrac{G^2\Sigma_t}{\gamma^2}\right)\hat{\mu}_t dt + \tfrac{G\Sigma_t}{\gamma^2} ds_t, \qquad \hat{\mu}_0 = a \\
d\Sigma_t &= \left(2F\Sigma_t + \sigma_f^2 - \tfrac{G^2\Sigma_t^2}{\gamma^2}\right) dt + \sigma_u^2 dQ_t, \quad \Sigma_0 = b
\end{aligned} \tag{A.4}$$

Furthermore, the first filtering equation can be written as

$$d\hat{\mu}_t = F\hat{\mu}_t dt + \frac{G^2\Sigma_t}{\gamma} d\hat{Z}_t$$

where $\hat{Z}_t$ *is the innovation process given by* $d\hat{Z}_t = \frac{1}{\gamma}(ds_t - \hat{\mu}_t dt) = \frac{1}{\gamma}(\mu_t - \hat{\mu}_t)dt + dZ_t$ *and it is one-dimensional Wiener process under the probability distribution of the firm independent of* dQ_t.

Finally, using the definition of uncertainty $\Omega_t \equiv \gamma\Sigma_t$, *and substituting* $F = 0$ *and* $G = 1$, *we obtain the filtering equations used in the text:*

$$d\hat{\mu}_t = \Omega_t d\hat{Z}_t, \qquad \hat{\mu}_0 = a \tag{A.5}$$

$$d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t, \qquad \Omega_0 = \frac{b}{\gamma} \tag{A.6}$$

Proof. The strategy of the proof has three steps, each established in a Lemma.

(I) We show that the solution $M_t \equiv [\mu_t, s_t]$ to the system of stochastic differential equations in (A.3), conditional on the history of Poisson shocks $\mathcal{Q}_t = \sigma\{Q_r | r \leq t\}$, follows a Gaussian process.

(II) $\mu_t|\mathcal{I}_t$ is Normal and can be obtained as the limit of a discrete sampling of observations;

(III) The recursive estimation formulas obtained with discrete sampling converge to (A.4).[9]

Now we elaborate on the three steps.

Lemma 1. *Let* $M_t \equiv [\mu_t, s_t]$ *be the solution to* (A.3) *and* $\mathcal{Q}_t = \sigma\{Q_r | r \leq t\}$. *Then* $M_t|\mathcal{Q}_t$ *is Normal.*

Proof. Fix a realization ω and let $N_t(\omega)$ be the quantity of jumps between 0 and t, which is a number known at t. Applying Picard iterative process to (A.3) and considering the initial conditions, we obtain the following sequences

$$\begin{aligned}
\mu_t^{k+1} &= \mu_0 + F\int_0^t \mu_\tau^k d\tau + \sigma_f W_t + \sigma_f \sum_{i=1}^{N_t(\omega)} u_i \\
s_t^{k+1} &= G\int_0^t \mu_\tau^k d\tau + \gamma Z_t
\end{aligned}$$

Assume that μ_t^0 is Normal. As an induction hypothesis, assume that $M_r^k|\mathcal{Q}_t \equiv [\mu_r^k, s_r^k|\mathcal{Q}_t]$ is Normal for all $r \leq t$. Note that (μ_0, W_r, Z_r) are Normal random variables independent of $\mathcal{Q}_t$; the term $\sum_{i=1}^{N_r(\omega)} u_i|\mathcal{Q}_t$ is Normal since it is a fixed sum of $N_r(\omega)$ Normal random variables; and finally, the term $\int_0^r \mu_\tau^k d\tau$ is a Riemann integral of Normal variables by the induction hypothesis. Given that the linear combination of Normals is Normal, then $M_r^{k+1}|\mathcal{Q}_t \equiv [\mu_r^{k+1}, s_r^{k+1}|\mathcal{Q}_t]$ is Normal as well for $r \leq t$. Therefore, for each $r \leq t$, we have a sequence of Normal random variables $\{M_r^k|\mathcal{Q}_t\}_{k=0}^\infty$.

To show Normality of $M_t|\mathcal{Q}_t$, notice that $M_r^k|\mathcal{Q}_t = M_r^k|\mathcal{Q}_r$ and $M_r^k|\mathcal{Q}_r$ converges in L^2 to M_r (see chapter 5 of Øksendal (2007)). Since the limit in L^2 of Normal variables is Normal, M_t is Normal. Therefore the solution to the system of stochastic differential equations, conditional to the history of Poisson shocks, i.e. $M_t|\mathcal{Q}_t$, is a Gaussian process. □

[9] In Section I of the Online Appendix, we derive additional details and a formal convergence proof.

Lemma 2. *The conditional distribution of the state $\mu_t|\mathcal{I}_t$ is Normal, $\mu_t|\mathcal{I}_t \sim \mathcal{N}\left(\mathbb{E}\left[\mu_t|\mathcal{I}_t\right], \mathbb{E}\left[\left(\mu_t - \mathbb{E}\left[\mu_t|\mathcal{I}_t\right]\right)^2|\mathcal{I}_t\right]\right)$, and the conditional mean and variance can be obtained as the limit of a discrete sampling of observations.*

Proof. Let $\Delta \equiv \frac{1}{2^n}$ and define an increasing sequence of σ-algebras $\{\mathcal{I}_t^n\}_{n=0}^{\infty}$ using the dyadic set as follows:

$$\mathcal{I}_t^n = \sigma\left\{s_r, Q_h : r \in \{0, \Delta, 2\Delta, 3\Delta, \dots\}, \ r \le t, \ h \in [0,t]\right\}$$

Let $M_t^n \equiv \mu_t|\mathcal{I}_t^n$ be the estimate at time t produced with discrete sampling. The following properties are true.

(i) For each n, M_t^n is a Normal random variable. By the previous Lemma $(\mu_t, s_{r_1}, s_{r_2}, \dots, s_{r_n})|\mathcal{Q}_t$ is Normal; by properties of Normals, M_t^n is also Normal.

(ii) For each n, M_t^n has finite variance. This is a direct implication of Normality.

(iii) Let $\mathcal{I}_t^{\infty} \equiv \sigma\{U_{n=1}^{\infty} I_t^n\}$ be the σ-algebra generated by the union of the discrete sampling information sets. For each t, M_t^n converges to some limit $M_t^{\infty} \equiv \mu_t|\mathcal{I}_t^{\infty}$ as $n \to \infty$. Since $\mathcal{I}_t^n$ is a increasing sequence of σ-algebras, by the Law of Iterated Expectations M_t^n is a martingale with finite variance, therefore it converges in L^2. Given that the limit of Normal random variables is Normal, the limit M_t^{∞} is a Normal random variable as well.

$$M_t^n \to_{L^2} M_t^{\infty} \ \sim \ \mathcal{N}(\mathbb{E}\left[\mu_t|\mathcal{I}_t^{\infty}\right], \mathbb{E}\left[\left(\mu_t - \mathbb{E}\left[\mu_t|\mathcal{I}_t\right]\right)^2|\mathcal{I}_t^{\infty}\right]$$

Since signals s_t are continuous (in particular left-continuous) and the dyadic set is dense in the interval $[0,t]$, the information set obtained as the limit of the discrete sampling is equal to the information set obtained with continuous sampling: $\mathcal{I}_t^{\infty} = \sigma\left\{s_h, Q_h : h \in [0,t]\right\}$. Therefore, the estimate obtained with the limit of discrete sampling converges (in L^2) to the estimate with continuos sampling (see Davis (1977) for more details in this topic).

$$M_t^{\infty} \to_{L^2} \mu_t|\mathcal{I}_t \ \sim \ \mathcal{N}\left(\mathbb{E}\left[\mu_t|\mathcal{I}_t\right], \mathbb{E}\left[\left(\mu_t - \mathbb{E}\left[\mu_t|\mathcal{I}_t\right]\right)^2|\mathcal{I}_t\right]\right)$$

□

Lemma 3. *Let $\Delta \equiv \frac{1}{2^n}$ and define $\mathcal{I}_t^{n,*}$ as the information set before measurement (used to construct predicted estimates)*

$$\mathcal{I}_t^{n,*} = \sigma\left\{s_{r-1}, Q_h \,| r \in \{0, \Delta, 2\Delta, 3\Delta, \dots\}, \ r \le t, \ h \in [0,t]\right\}$$

and define $\hat{\mu}_t^n = \mathbb{E}[\mu_t|\mathcal{I}_t^{n,*}]$ and $\Sigma_t^n = \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|\mathcal{I}_t^{n,*}]$. Then the laws of motion of $\{\hat{\mu}_t^n, \Sigma_t^n\}$ converge weakly to the solution of (A.4), namely the laws of motion for $\{\hat{\mu}_t, \Sigma_t\}$, where $\hat{\mu}_t \equiv \mathbb{E}[\mu_t|\mathcal{I}_t]$ and $\Sigma_t \equiv \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|\mathcal{I}_t]$.

Proof. Before we derive the processes for the estimate and its conditional variance, an explanation of why we use the information set $\mathcal{I}_t^{n,*}$ instead of $\mathcal{I}_t^n$ is due. The reason is convenience, as the first information set produces independent recursive formulas for the predicted estimate $\mu_t|\sigma\{U_{i=1}^{\infty} I_t^{n,*}\}$ and it is easier to show its convergence. Let us show that the union of information sets are equal, i.e. $\sigma\{U_{i=1}^{\infty} I_t^n\} = \sigma\{U_{i=1}^{\infty} I_t^{n,*}\}$, and thus the way we construct the limit is innocuous. Trivially, we have that $\sigma\{U_{i=1}^{\infty} I_t^{n,*}\} \subset \sigma\{U_{i=1}^{\infty} I_t^n\}$. For the reverse to be true $\sigma\{U_{i=1}^{\infty} I_t^n\} \subset \sigma\{U_{i=1}^{\infty} I_t^{n,*}\}$, it is sufficient to show that signals s are continuous, since left-continuous filtrations of continuous process are always continuous. To show that signals are continuous, notice that they can be written as $s_t = \int_0^t \mu_s ds + \gamma Z_t$, which is an integral of a finite set of discontinuities plus a Wiener process, and thus they are continuous.

Now let us derive the laws of motion. Considering an interval Δ, then the processes in (A.3) can be written as

$$\begin{aligned}
\mu_t &= \mu_{t-\Delta} + F\int_{t-\Delta}^{t} \mu_\tau d\tau + \sqrt{\Delta\sigma_f^2}\epsilon_t + \sigma_u u_t(Q_t - Q_{t-\Delta}), \qquad \mu_{-\Delta} \sim \mathcal{N}(\hat{\mu}_\Delta, \Sigma_\Delta) \\
s_t &= s_{t-\Delta} + G\int_{t-\Delta}^{t} \mu_\tau d\tau + \sqrt{\Delta\gamma^2}\eta_t, \qquad s_0 = 0 \\
(Q_t - Q_{t-\Delta}) &\sim_{i.i.d} \begin{cases} 1 & \text{with probability } 1 - e^{-\lambda\Delta} - o(\Delta^2) \\ 0 & \text{with probability } e^{-\lambda\Delta} - o(\Delta^2) \\ > 1 & \text{with probability } o(\Delta^2) \end{cases} \\
\epsilon_t, \eta_t, u_t &\sim_{i.i.d} \mathcal{N}(0,1)
\end{aligned}$$

First order approximations of the integral yield $\int_{t-\Delta}^{t} \mu_\tau d\tau = \mu_{t-\Delta}\Delta + \xi_t = \mu_t\Delta + \tilde{\xi}_t$, where ξ_t and $\tilde{\xi}_t$ are Normal random variables conditional on $\mathcal{Q}_t$, with $\mathbb{E}[\xi_t] = o(\Delta^2)$, $\mathbb{E}[\xi_t^2] = o(\Delta^2)$, $\mathbb{E}[\tilde{\xi}_t] = o(\Delta^2)$ and $\mathbb{E}[\tilde{\xi}_t] = o(\Delta^2)$. Substituting these approximations above, we can express the laws of motion for μ, s as follows:

$$\begin{aligned}
\mu_t &= (1 + F\Delta)\mu_{t-\Delta} + \sqrt{\Delta\sigma_f^2}\epsilon_t + \sigma_u u_t(Q_t - Q_{t-\Delta}) + o(\Delta^2) \\
s_t &= s_{t-\Delta} + G\Delta\mu_t + \sqrt{\Delta\gamma^2}\eta_t + o(\Delta^2)
\end{aligned}$$

Since the model is Gaussian, we use the Kalman Filter to estimate the conditional mean $\hat{\mu}_t^n = \mathbb{E}[\mu_t | \mathcal{I}_t^{n,*}]$ and variance $\Sigma_t^n = \mathbb{E}[(\mu_t - \hat{\mu}_t)^2 | \mathcal{I}_t^{n,*}]$. The recursive formulas are

$$\begin{aligned}
\hat{\mu}_{t+\Delta}^n &= (1+\Delta F)\,\hat{\mu}_t^n + K_t^n\,(s_t - s_{t-\Delta} - \Delta G(1+\Delta F)\hat{\mu}_t^n) + o(\Delta^2) \\
\Sigma_{t+\Delta}^n &= (1+\Delta F)^2 \frac{\Sigma_t^n \gamma^2}{\Sigma_t^n G^2 \Delta + \gamma^2} + \sigma_f^2 \Delta + (Q_{t+\Delta} - Q_t)\sigma_u^2 + o(\Delta^2) \\
K_t^n &= (1+\Delta F)\,\frac{\Sigma_t^n G}{\Sigma_t^n G^2 \Delta + \gamma^2}
\end{aligned}$$

Notice that since u_t has mean zero, the known arrival of a Poisson shock does not affect the estimate. However, it does affect the variance by adding a shock of size σ_u^2. Rearranging and doing some algebra, the previous system can be written as

$$\begin{aligned}
\hat{\mu}_{t+\Delta}^n - \hat{\mu}_t^n &= \left(F - G\varphi^I(\Delta)\right)\hat{\mu}_t \Delta + \varphi^I(\Delta)\,(s_t - s_{t-\Delta}) + o(\Delta^2), & \varphi^I(\Delta) &\equiv \frac{\Sigma_t^n G}{\Sigma_t^2 G^2 \Delta + \gamma^2} \\
\Sigma_{t+\Delta}^n - \Sigma_t^n &= \left(\varphi^{II}(\Delta) + \sigma_f^2\right)\Delta + (Q_{t+\Delta} - Q_t)\sigma_u^2 + o(\Delta^2), & \varphi^{II}(\Delta) &\equiv \left[\frac{\gamma^2(2F + F^2\Delta) - G^2\Sigma_t^n}{\Sigma_t^n G^2 \Delta + \gamma^2}\right]\Sigma_t^n
\end{aligned}$$

Taking the limit as $n \to \infty$ (or $\Delta \to 0$), we see that $\varphi^I(\Delta) \to \frac{\Sigma_t G}{\gamma^2}$ and $\varphi^{II}(\Delta) \to 2F\Sigma_t - \frac{G^2\Sigma_t^2}{\gamma^2}$, which yield exactly the same laws of motion that can be obtained with the continuous time Kalman-Bucy filter. Therefore, the laws of motion obtained with discrete sampling are locally consistent with the continuous time filtering equations in (A.4) (see Section I of the Online Appendix for more details, where we follow closely Theorem 1.1, Chapter 10 of Kushner and Dupuis (2001).). □

To conclude the proof, use the structure of the signal to rewrite the law of motion in innovation representation as

$$d\hat{\mu}_t = F\hat{\mu}_t dt + \frac{G\Sigma_t}{\gamma}\left(\frac{G}{\gamma}(\mu_t - \hat{\mu}_t)\,dt + dZ_t\right) = F\hat{\mu}_t dt + \frac{G\Sigma_t}{\gamma} d\hat{Z}_t \tag{A.7}$$

where $d\hat{Z}_t \equiv \frac{G}{\gamma}(\mu_t - \hat{\mu}_t)\,dt + dZ_t$ is the innovation process. We now show $d\hat{Z}_t$ is a Wiener process. Applying the law of iterated expectations:

$$\mathbb{E}[(\mu_t - \hat{\mu}_t)|\sigma\{\hat{\mu}_s : s \le t\}] = \mathbb{E}\left[\mathbb{E}\left[(\mu_t - \hat{\mu}_t)|\,\mathcal{I}_t\right]|\sigma\{\hat{\mu}_s : s \le t\}\right] = \mathbb{E}[(\hat{\mu}_t - \hat{\mu}_t)|\sigma\{\hat{\mu}_s : s \le t\}] = 0$$

Since $\mathbb{E}[(\mu_t - \hat{\mu}_t)|\sigma\{\hat{\mu}_s : s \le t\}] = 0\ \forall t$ and dZ_t is a Wiener process, we apply corollary 8.4.5 of Øksendal (2007) and conclude that $d\hat{Z}_t$ is a Weiner process as well.

□

Proposition 2 (**Stopping time problem**). *Let $(\hat{\mu}_0, \Omega_0)$ be the firm's current state immediately after the last markup adjustment. Also let $\bar{\theta} = \frac{\theta}{B}$ be the normalized menu cost. Then the optimal stopping time and reset markup gap (τ, μ') solve the following problem:*

$$V(\hat{\mu}_0, \Omega_0) = \max_{\tau} \mathbb{E}\left[\int_0^\tau -e^{-rs}\hat{\mu}_s^2 ds + e^{-r\tau}\Big(-\bar{\theta} + \max_{\mu'} V(\mu', \Omega_\tau)\Big)\Big|\mathcal{I}_0\right] \tag{9}$$

subject to the filtering equations in Proposition 1.

Proof. Let $\{\tau_i\}_{i=1}^\infty$ be the series of dates where the firm adjusts her markup gap and $\{\mu_i\}_{i=1}^\infty$ the series of reset markup gaps. Given an initial condition μ_0 and a law of motion for the markup gaps, the sequential problem of the firm is expressed as follows:

$$\max_{\{\mu_{\tau_i}, \tau_i\}_{i=1}^\infty} \mathbb{E}\left[\sum_{i=0}^\infty e^{-r\tau_{i+1}}\left(-\bar{\theta} - \int_{\tau_i}^{\tau_{i+1}} e^{-r(t-\tau_{i+1})}\mu_t^2 dt\right)\right] \tag{A.8}$$

Using the definition of variance, we can write the condition expectation of the markup gap at time t as:

$$\mathbb{E}[\mu_t^2|\mathcal{I}_t] = \mathbb{E}[\mu_t|\mathcal{I}_t]^2 + \mathbb{V}[\mu_t|\mathcal{I}_t] = \hat{\mu}_t^2 + \mathbb{V}[\mu_t|\mathcal{I}_t] = \hat{\mu}_t^2 + (\sigma_f^2 + \lambda\sigma_u^2)t = \hat{\mu}_t^2 + \Omega^{*2}t$$

where in the last equality we use the definition of fundamental uncertainty Ω^*. Use the Law of Iterated Expectations in (A.8) to take expectation given the information set at time t. Use the decomposition above to write the problem in terms of estimates:

$$\mathbb{E}\left[\sum_{i=0}^\infty e^{-r\tau_{i+1}}\left(-\bar{\theta} - \int_{\tau_i}^{\tau_{i+1}} e^{-r(t-\tau_{i+1})}\mathbb{E}\left[\mu_t^2\Big|\mathcal{I}_t\right]dt\right)\right]$$

$$\mathbb{E}\left[\sum_{i=0}^\infty e^{-r\tau_{i+1}}\left(-\bar{\theta} - \int_{\tau_i}^{\tau_{i+1}} e^{-r(t-\tau_{i+1})}(\hat{\mu}_t^2 + \Omega^{*2}t)dt\right)\right]$$

$$\mathbb{E}\left[\sum_{i=0}^\infty e^{-r\tau_{i+1}}\left(-\bar{\theta} - \int_{\tau_i}^{\tau_{i+1}} e^{-r(t-\tau_{i+1})}\hat{\mu}_t^2 dt\right)\right] - \underbrace{\Omega^{*2}\mathbb{E}\left[\sum_{i=0}^\infty \int_{\tau_i}^{\tau_{i+1}} te^{-rt}dt\right]}_{sunk\ cost}$$

The last term in the previous expression is a constant number, and it arises from the fact that the firm will never learn the true realization of the markup gap. It is considered a sunk cost in the firm's problem since she cannot take any action to alter its value; therefore, we can ignore it from her problem. To compute its value, note that the term inside the expectation is equal to:

$$\sum_{i=0}^\infty \int_{\tau_i}^{\tau_{i+1}} te^{-rt}dt \quad = \quad \sum_{i=0}^\infty \left[\frac{e^{-r\tau_i}(1+r\tau_i) - e^{-r\tau_{i+1}}(1+r\tau_{i+1})}{r^2}\right] = \frac{e^{-r\tau_0}(1+r\tau_0)}{r^2}$$

where the sum is telescopic and all terms except the first cancel out. Therefore, the sunk cost term becomes:

$$-\Omega^{*2}\mathbb{E}\left[\frac{e^{-r\tau_0}(1+r\tau_0)}{r^2}\right] < \infty$$

Using the previous results, the sequential problem in (A.8) can be written in terms of estimates instead of the true realizations:

$$\max_{\{\mu_{\tau_i}, \tau_i\}_{i=1}^\infty} \mathbb{E}\left[\sum_{i=0}^\infty e^{-r\tau_{i+1}}\left(-\bar{\theta} - \int_{\tau_i}^{\tau_{i+1}} e^{-r(t-\tau_{i+1})}\hat{\mu}_t^2 dt\right)\right]$$

Given the stationarity of the problem and the stochastic processes, we apply the Principle of Optimality to the sequential problem (see equation 7.2 in Stokey (2009)) and express it as a sequence of stopping time problems with state $(\hat{\mu}_0, \Omega_0)$:

$$V(\hat{\mu}_0, \Omega_0) = \max_{\tau} \mathbb{E}\left[\int_0^\tau -e^{-rt}\hat{\mu}_t^2 dt + e^{-r\tau}[-\bar{\theta} + \max_{\mu'} V(\mu', \Omega_\tau)]\right]$$

subject to the filtering equations. Here τ is the stopping time associated with the optimal decision. □

Proposition 3 (**HJB Equation, Value Matching and Smooth Pasting**). *Let $\phi : R \times R^+ \to R$ be a function and let ϕ_x denote the derivative of ϕ with respect to x. Assume ϕ satisfies the following conditions:*

1. *For all states in the interior of the inaction region $\mathcal{R}^o$, ϕ solves the* <u>*Hamilton-Jacobi-Bellman*</u> *(HJB) equation:*

$$r\phi(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 - \Omega^2}{\gamma}\right)\phi_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2}\phi_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda\left[\phi\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - \phi(\hat{\mu}, \Omega)\right] \tag{10}$$

2. *At the border of the inaction region $\partial\mathcal{R}$, ϕ satisfies the* <u>*value matching condition*</u>*, which sets the value of adjusting equal to the value of not adjusting:*

$$\phi(0, \Omega) - \bar{\theta} = \phi(\bar{\mu}(\Omega), \Omega) \tag{11}$$

3. *At the border of the inaction region $\partial\mathcal{R}$, ϕ satisfies* <u>*two smooth pasting conditions*</u>*, one for each state:*

$$\phi_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega) = 0, \quad \phi_\Omega(\bar{\mu}(\Omega), \Omega) = \phi_\Omega(0, \Omega) \tag{12}$$

Then ϕ is the value function $\phi = V$ and $\tau = \inf\{t > 0 : \phi(0, \Omega_t) - \theta > \phi(\hat{\mu}_t, \Omega_t)\}$ is the optimal stopping time.

Proof. Start from the recursive representation of the value function as a stopping time problem derived in Proposition 2.

$$\begin{aligned} V(\hat{\mu}_0, \Omega_0) &= \max_\tau \mathbb{E}\left[\int_0^\tau -e^{-rt}\hat{\mu}_t^2 dt + e^{-r\tau}[-\bar{\theta} + \max_{\mu'} V(\mu', \Omega_\tau)]\right] \\ d\hat{\mu}_t &= \Omega_t d\hat{Z}_t, \qquad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma}dt + \frac{\sigma_u^2}{\gamma}dQ_t \end{aligned}$$

To obtain the HJB, consider the value function inside the continuation region. Then for a small interval dt we can write:

$$\begin{aligned} V(\hat{\mu}_t, \Omega_t) &= -\hat{\mu}_t^2 dt + e^{-rdt}\mathbb{E}[V(\hat{\mu}_{t+dt}, \Omega_{t+dt})] \\ (1 - e^{-rdt})V(\hat{\mu}_t, \Omega_t) &= -\hat{\mu}_t^2 dt + e^{-rdt}\mathbb{E}[V(\hat{\mu}_{t+dt}, \Omega_{t+dt}) - V(\hat{\mu}_t, \Omega_t)] \\ rV(\hat{\mu}_t, \Omega_t)dt &= -\hat{\mu}_t^2 dt + (1 - rdt)\mathbb{E}[V(\hat{\mu}_{t+dt}, \Omega_{t+dt}) - V(\hat{\mu}_t, \Omega_t)] \\ rV(\hat{\mu}_t, \Omega_t) &= -\hat{\mu}_t^2 + \lim_{dt \downarrow 0}(1 - rdt)\frac{\mathbb{E}[V(\hat{\mu}_{t+dt}, \Omega_{t+dt}) - V(\hat{\mu}_t, \Omega_t)]}{dt} \\ rV(\hat{\mu}_t, \Omega_t) &= -\hat{\mu}_t^2 + \mathcal{A}V(\hat{\mu}_t, \Omega_t) \end{aligned}$$

where in the second line we have subtracted $e^{-rdt}V(\hat{\mu}_t, \Omega_t)$ from both sides, in the third line we have approximated e^{-rdt} with $1 - rdt$, in the fourth line we divide by dt and then take the limit $dt \to 0$, and finally in the fifth line we recognized the definition of the generator. Substituting the generator $\mathcal{A}$ from (A.1) we obtain the HJB equation:

$$rV(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 - \Omega^2}{\gamma}\right)V_\Omega(\hat{\mu}, \Omega) + \frac{1}{2}\Omega^2 V_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda\left[V\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - V(\hat{\mu}, \Omega)\right]$$

The value matching condition that sets equal the value of adjusting and not adjusting at the border:

$$V(\bar{\mu}(\Omega), \Omega) = V(0, \Omega) - \bar{\theta}$$

We apply Theorem 2.2 of Øksendal and Sulem (2010) and impose two smooth pasting conditions, one for each state,

$$V_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega) = 0, \qquad V_\Omega(\bar{\mu}(\Omega), \Omega) = V_\Omega(0, \Omega)$$

Section B of the Online Appendix verifies that the conditions in that Theorem hold in our problem; and Section C.3 verifies numerically that the smooth pasting conditions for $\hat{\mu}$ and Ω are valid. ☐

Proposition 4 (Inaction region). *For r and $\bar{\theta}$ be small, the border of the inaction region is approximated by*

$$\bar{\mu}(\Omega) = \left(\frac{6\bar{\theta}\Omega^2}{1+\mathcal{L}^{\bar{\mu}}(\Omega)}\right)^{1/4}, \quad \text{with} \quad \mathcal{L}^{\bar{\mu}}(\Omega) = \left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{\Omega}{\Omega^*}-1\right) \tag{13}$$

The elasticity of $\bar{\mu}(\Omega)$ with respect to Ω is equal to

$$\mathcal{E}(\Omega) \equiv \frac{1}{2} - \left(\frac{1}{6}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\frac{\Omega}{\Omega^*} \tag{14}$$

Lastly, the reset markup gap is equal to $\hat{\mu}' = 0$.

Proof. The plan for the proof is as follows. Following Álvarez, Lippi and Paciello (2011) we use Taylor approximations to the value function and optimality conditions to characterize the border of the inaction region. We first obtain an inaction region that depends on derivatives of the value function. This derivatives introduce a novel term – which we label learning component – that does not appear in inaction regions derived in perfect information settings. We then approximate this learning component around fundamental uncertainty Ω^*. With this approximation, we obtain an expression for the inaction region that depends only on the uncertainty level and parameters. Lastly, we show that the elasticity of the inaction region with respect to uncertainty is lower than unity; and that the reset markup gap is equal to zero.

1. **Optimality conditions:** The optimality conditions of the problem are given by:

$$rV(\hat{\mu},\Omega) = -\hat{\mu}^2 + \lambda\left[V\left(\hat{\mu},\Omega+\frac{\sigma_u^2}{\gamma}\right) - V(\hat{\mu},\Omega)\right] + \frac{\sigma_f^2-\Omega^2}{\gamma}V_\Omega(\hat{\mu},\Omega) + \frac{\Omega^2}{2}V_{\hat{\mu}^2}(\hat{\mu},\Omega) \tag{A.9}$$

$$V(\bar{\mu}(\Omega),\Omega) = V(0,\Omega) - \bar{\theta} \tag{A.10}$$

$$V_\mu(\bar{\mu}(\Omega),\Omega) = 0 \tag{A.11}$$

$$V_\Omega(\bar{\mu}(\Omega),\Omega) = V_\Omega(0,\Omega) \tag{A.12}$$

2. **Taylor approximation of V and value matching**. For a given level of uncertainty Ω, we do a 4^{th} order Taylor expansion on the first argument of V around zero:

$$V(\hat{\mu},\Omega) = V(0,\Omega) + \frac{V_{\hat{\mu}^2}(0,\Omega)}{2!}\hat{\mu}^2 + \frac{V_{\hat{\mu}^4}(0,\Omega)}{4!}\hat{\mu}^4$$

Odd terms do not appear due to the symmetry of the value function around 0. Evaluating at the border and combining with the value matching condition (A.10) we obtain:

$$-\bar{\theta} = V_{\hat{\mu}^2}(0,\Omega)\frac{\bar{\mu}(\Omega)^2}{2} + V_{\hat{\mu}^4}(0,\Omega)\frac{\bar{\mu}(\Omega)^4}{24} \tag{A.13}$$

3. **Taylor approximation of V_μ and smooth pasting**. For a given level of uncertainty Ω, we do a 3^{rd} order Taylor expansion on the first argument of V_μ around zero:

$$V_\mu(\hat{\mu},\Omega) = V_{\hat{\mu}^2}(0,\Omega)\hat{\mu} + \frac{V_{\hat{\mu}^4}(0,\Omega)}{3!}\hat{\mu}^3$$

Again the odd derivatives are zero. Evaluate at the border, multiply both sides by $\frac{\bar{\mu}(\Omega)}{2}$ and combine with the smooth pasting condition (A.11) to obtain:

$$0 = V_{\hat{\mu}^2}(0,\Omega)\frac{\bar{\mu}(\Omega)^2}{2} + V_{\hat{\mu}^4}(0,\Omega)\frac{\bar{\mu}(\Omega)^4}{12} \tag{A.14}$$

4. **Inaction border (as a function of V)**: Combine the relationships between the 2nd and 4th derivatives of V in (A.13) and (A.14):

$$\bar{\theta} = \bar{\mu}(\Omega)^4\frac{V_{\hat{\mu}^4}(0,\Omega)}{24} = -\bar{\mu}(\Omega)^2\frac{V_{\hat{\mu}^2}(0,\Omega)}{4} \tag{A.15}$$

From the previous equality, we obtain an expression for the border of inaction as a function of $V_{\hat{\mu}^4}$:

$$\bar{\mu}(\Omega) = \left(\frac{24\bar{\theta}}{V_{\hat{\mu}^4}(0,\Omega)}\right)^{1/4} \tag{A.16}$$

5. **Definition of learning effect $\mathcal{L}^{\bar{\mu}}(\Omega)$**: We want to further characterize $V_{\hat{\mu}^4}(0,\Omega)$. Taking second derivatives of the HBJ in (A.9) with respect to $\hat{\mu}$:

$$rV_{\hat{\mu}^2}(\hat{\mu},\Omega) = -2 + \lambda\left(V_{\hat{\mu}^2}\left(\hat{\mu},\Omega+\frac{\sigma_u^2}{\gamma}\right) - V_{\hat{\mu}^2}(\hat{\mu},\Omega)\right) + \frac{\sigma_f^2-\Omega^2}{\gamma}V_{\hat{\mu}^2\Omega}(\hat{\mu},\Omega) + \frac{\Omega^2}{2}V_{\hat{\mu}^4}(\hat{\mu},\Omega)$$

Now we use a Taylor approximation of the second argument of $V_{\hat{\mu}^2}\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right)$ around Ω:

$$V_{\hat{\mu}^2}\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) = V_{\hat{\mu}^2}\left(\hat{\mu}, \Omega\right) + V_{\hat{\mu}^2\Omega}\left(\hat{\mu}, \Omega\right)\frac{\sigma_u^2}{\gamma}$$

Substitute back this expression and use the definition of fundamental uncertainty $\Omega^{*2} \equiv \sigma_f^2 + \lambda\sigma_u^2$ to get:

$$rV_{\hat{\mu}^2}(\hat{\mu}, \Omega) = -2 - \frac{\Omega^2 - \Omega^{*2}}{\gamma} V_{\hat{\mu}^2\Omega}(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} V_{\hat{\mu}^4}(\hat{\mu}, \Omega)$$

Lastly, taking the limit $r \to 0$, evaluating at $\hat{\mu} = 0$, and rearranging:

$$V_{\hat{\mu}^4}(0, \Omega) = \frac{4}{\Omega^2}\left(1 + \frac{\Omega^2 - \Omega^{*2}}{\gamma}\frac{V_{\hat{\mu}^2\Omega}(0, \Omega)}{2}\right) = \frac{4}{\Omega^2}\left(1 + \mathcal{L}^{\bar{\mu}}(\Omega)\right) \tag{A.17}$$

where we define the learning component as

$$\mathcal{L}^{\bar{\mu}}(\Omega) \equiv \frac{\Omega^2 - \Omega^{*2}}{\gamma}\frac{V_{\hat{\mu}^2\Omega}(0, \Omega)}{2} \tag{A.18}$$

With this expression for $V_{\hat{\mu}^4}(0, \Omega)$, the border of the inaction region in (A.16) changes to:

$$\bar{\mu}(\Omega) = \left(\frac{6\bar{\theta}\Omega^2}{1 + \mathcal{L}^{\bar{\mu}}(\Omega)}\right)^{1/4} \tag{A.19}$$

Note that if $\Omega = \Omega^*$, then $\mathcal{L}^{\bar{\mu}}(\Omega) = 0$ and the formula for the inaction region collapses to 4-th root formula in Dixit (1991) and Álvarez, Lippi and Paciello (2011), where Ω takes the place of σ_f.

6. **Approximate learning component $\mathcal{L}^{\bar{\mu}}(\Omega)$ around Ω^*.** Define $\Gamma(\Omega) \equiv \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}$. To characterize it, first use the equivalence between the 2nd and 4th derivatives in (A.15), then substitute the expressions for $\mathcal{L}^{\bar{\mu}}(\Omega)$ in (A.17) and $\bar{\mu}(\Omega)$ in (A.18), and then simplify:

$$\Gamma(\Omega) \equiv \frac{\partial}{\partial\Omega}\left[\frac{V_{\hat{\mu}^2}(0, \Omega)}{2}\right] = \frac{\partial}{\partial\Omega}\left[-\frac{V_{\hat{\mu}^4}(0, \Omega)}{12}\bar{\mu}(\Omega)^2\right] = \frac{\partial}{\partial\Omega}\left[-\frac{(\frac{2}{3}\bar{\theta})^{1/2}}{\Omega}\left(1 + \mathcal{L}^{\bar{\mu}}(\Omega)\right)^{1/2}\right]$$

Using the definition of $\Gamma(\Omega)$ write the previous equation recursively as:

$$\Gamma(\Omega) = \frac{\partial}{\partial\Omega}\left[-\frac{(\frac{2}{3}\bar{\theta})^{1/2}}{\Omega}\left(1 + \Gamma(\Omega)\frac{\Omega^2 - \Omega^{*2}}{\gamma}\right)^{1/2}\right]$$

A first order Taylor approximation of $\mathcal{L}^{\bar{\mu}}(\Omega)$ around Ω^* yields:

$$\mathcal{L}^{\bar{\mu}}(\Omega) = \mathcal{L}^{\bar{\mu}}(\Omega^*) + \mathcal{L}^{\bar{\mu}}_{\Omega}(\Omega^*)(\Omega - \Omega^*) = 2\Omega^*\Gamma(\Omega^*)\frac{\Omega - \Omega^*}{\gamma} = \left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{\Omega}{\Omega^*} - 1\right)$$

where we have used the following equalities: $\mathcal{L}^{\bar{\mu}}(\Omega^*) = 0$, $\mathcal{L}^{\bar{\mu}}_{\Omega}(\Omega^*) = 2\frac{\Omega^*}{\gamma}\Gamma(\Omega^*)$, and $\Gamma(\Omega^*) = \frac{(\frac{2}{3}\bar{\theta})^{1/2}}{\Omega^{*2}}$. Substituting back into the border, we get the final approximation:

$$\bar{\mu}(\Omega) = (6\bar{\theta}\Omega^2)^{1/4}\left[1 + \left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{\Omega}{\Omega^*} - 1\right)\right]^{-1/4} \tag{A.20}$$

7. **Elasticity:** Now we compute the elasticity of the border to uncertainty $\mathcal{E} \equiv \frac{\partial \ln\bar{\mu}(\Omega)}{\partial \ln\Omega}$.
Applying logs to (A.20) we obtain:

$$\ln\bar{\mu}(\Omega) \propto \frac{1}{2}\ln\Omega - \frac{1}{4}\ln\left[1 + \left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{\Omega}{\Omega^*} - 1\right)\right]$$

Our parametric assumptions of small menu costs $\bar{\theta}$ and large signal noise γ make the quantity $\left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}$ very small, therefore, we use $\ln(1 + x) \approx x$ for x small to get:

$$\ln\bar{\mu}(\Omega) \propto \frac{1}{2}\ln\Omega - \frac{1}{4}\left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{e^{\ln\Omega}}{\Omega^*} - 1\right)$$

Taking the derivatives, we obtain the elasticity:

$$\mathcal{E} \equiv \frac{1}{2} - \left(\frac{1}{6}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\frac{\Omega}{\Omega^*} \tag{A.21}$$

Clearly, $\mathcal{E} < 1$. In fact, since Ω is bounded below by σ_f, the highest value for the elasticity is $\mathcal{E} = \frac{1}{2} - \left(\frac{1}{6}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\frac{\sigma_f}{\Omega^*} < \frac{1}{2}$.

8. **Smooth pasting condition for** Ω: Lastly, we show that the smooth pasting condition for Ω is implied by other conditions. First, recall from (A.15) that

$$-\bar{\theta} = \bar{\mu}(\Omega)^2 \frac{V_{\hat{\mu}^2}(0,\Omega)}{4}$$

Write the RHS as:

$$\bar{\mu}(\Omega)^2 \frac{V_{\hat{\mu}^2}(0,\Omega)}{2} - \bar{\mu}(\Omega)^2 \frac{V_{\hat{\mu}^2}(0,\Omega)}{4} = \bar{\mu}(\Omega)^2 \frac{V_{\hat{\mu}^2}(0,\Omega)}{2!} + \bar{\mu}(\Omega)^4 \frac{V_{\hat{\mu}^4}(0,\Omega)}{4!} = V(\bar{\mu}(\Omega),\Omega) - V(0,\Omega)$$

where in the first equality we have substituted the equality in (A.15), and in the second equality we have used the HJB in (A.9) evaluated at $\bar{\mu}$. Summarizing, we have that

$$-\bar{\theta} = V(\bar{\mu}(\Omega),\Omega) - V(0,\Omega)$$

Finally, take derivative with respect to Ω on both sides and obtain the smooth pasting condition for Ω in (A.12).

$$0 = V_\Omega(\bar{\mu}(\Omega),\Omega) - V_\Omega(0,\Omega)$$

9. **The reset markup estimate is** $\hat{\mu}' = 0$**:** If we show that, V is symmetric at $\hat{\mu} = 0$ for any Ω, then V attains its critical value at zero as well. This can be seen using the definition of symmetric derivative. Note that the value of action is completely independent of the current state (and thus it is symmetric), since the instantaneous profit and future conditional expectations depend only on the new optimized estimate $\hat{\mu}'$, not the current one. Now let us consider the value of inaction:

$$-\hat{\mu}^2 dt + e^{-rdt}\mathbb{E}_t[V(\hat{\mu}',\Omega')|\hat{\mu}]$$

Since the instantaneous return is clearly symmetric, we are left to show that the conditional expectation of V is symmetric around zero.

$$\mathbb{E}_t[V(\hat{\mu}',\Omega')|\hat{\mu}] = \mathbb{E}_t[V(\hat{\mu}',\Omega')| - \hat{\mu}]$$

Since the stochastic process is symmetric around $\hat{\mu}$, the expectation of V using the distribution centered at $-\hat{\mu}$ or centered at $-\hat{\mu}$ gives the same expected value. Therefore, the value of inaction is symmetric around zero. In conclusion, the value function is symmetric around zero and thus, conditional on adjustment, the policy is to change the price so that the markup gap estimate goest to zero, i.e.

□

Proposition 5 (**Conditional Expected Time**). *Let r and $\bar{\theta}$ be small. The expected time for the next price change conditional on the state, denoted by $\mathbb{E}[\tau \big| \hat{\mu}, \Omega]$, is approximated as:*

$$\mathbb{E}[\tau \big| \hat{\mu}, \Omega] = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2} \left(1 + \mathcal{L}^{\tau}(\Omega)\right) \quad \textit{where} \quad \mathcal{L}^{\tau}(\Omega) \equiv 2\left(\frac{\Omega}{\Omega^*} - 1\right)\left(1 - \mathcal{E}(\Omega^*)\right)\left(\frac{\gamma (24\bar{\theta})^{1/2}}{\gamma + (24\bar{\theta})^{1/2}}\right) \tag{15}$$

If the elasticity of the inaction region with respect to uncertainty is lower than unity and signal noise is large, then the expected time between price changes (i.e. $\mathbb{E}[\tau \big| 0, \Omega]$) is a decreasing and convex function of uncertainty.

Proof. Let $T(\hat{\mu}, \Omega)$ denote the expected time for the next price change given the current state, i.e. $\mathbb{E}[\tau \big| \hat{\mu}, \Omega]$. The proof consists of four steps. First, we establish the HJB equation for $T(\hat{\mu}, \Omega)$ and its corresponding border condition. We apply a first order approximation to the HJB equation on the second state to compute the value with uncertainty jump. Second, we do a second order Taylor approximation of $T(\hat{\mu}, \Omega)$ around $(0, \Omega)$, and substitute both the HJB and the border condition into this approximation. This delivers an expression for the expected time that depends on two multiplicative terms: (i) the distance between the markup gap estimate and the border of the inaction region, normalized by uncertainty; and (ii) a term that measures the effect of uncertainty changes into the expected time. Third, we approximate term (ii) around fundamental uncertainty Ω^*. Lastly, we show that if $\mathcal{E} < 1$, then time for between price adjustments $T(0, \Omega)$ is decreasing in uncertainty.

1. **HJB equation, jump approximation, and border condition**. Consider a small interval dt. Then $T(\hat{\mu}, \Omega)$ can be written recursively as:

$$T(\hat{\mu}_t, \Omega_t) = 1dt + \mathbb{E}\left[T(\hat{\mu}_{t+dt}, \Omega_{t+dt})\right]$$

Passing T to the right hand side, dividing by dt and taking the limit $dt \to 0$:

$$0 = 1 + \lim_{dt \downarrow 0} \frac{\mathbb{E}\left[T(\hat{\mu}_{t+dt}, \Omega_{t+dt}) - T(\hat{\mu}_t, \Omega_t)\right]}{dt}$$

Recognizing the definition of the generator, we obtain the following HJB equation:

$$0 = 1 + \mathcal{A}T(\hat{\mu}, \Omega)$$

Substituting the infinitesimal generator $\mathcal{A}$ from (A.1) we obtain:

$$0 = 1 + \lambda\left[T\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - T(\hat{\mu}, \Omega)\right] + \frac{(\sigma_f^2 - \Omega^2)}{\gamma} T_{\Omega}(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} T_{\hat{\mu}^2}(\hat{\mu}, \Omega)$$

We approximate the uncertainty jump with a linear approximation to the second state:

$$T\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) \approx T(\hat{\mu}, \Omega) + \frac{\sigma_u^2}{\gamma} T_{\Omega}(\hat{\mu}, \Omega)$$

Substituting the approximation and using the definition of fundamental uncertainty Ω^*, we obtain:

$$0 = 1 + \frac{\Omega^{*2} - \Omega^2}{\gamma} T_{\Omega}(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} T_{\hat{\mu}^2}(\hat{\mu}, \Omega) \tag{A.22}$$

The border condition states that at the border of action, the expected time is equal to zero:

$$T(\bar{\mu}(\Omega), \Omega) = 0 \tag{A.23}$$

2. **Approximation of $T(\hat{\mu}, \Omega)$.** A second order Taylor approximation of $T(\hat{\mu}, \Omega)$ in the first state around $\hat{\mu} = 0$ yields:

$$T(\hat{\mu}, \Omega) = T(0, \Omega) + \frac{T_{\hat{\mu}^2}(0, \Omega)}{2} \hat{\mu}^2 \tag{A.24}$$

- To compute $T(0, \Omega)$, we evaluate (A.24) at $(\bar{\mu}(\Omega), \Omega)$ and use the border condition in (A.23):

$$T(0, \Omega) = -\frac{T_{\hat{\mu}^2}(0, \Omega)}{2} \bar{\mu}(\Omega)^2$$

- To compute $T_{\hat{\mu}^2}(0, \Omega)/2$, we evaluate the HJB in (A.22) at $(0, \Omega)$ and solve for it:

$$\frac{T_{\hat{\mu}^2}(0, \Omega)}{2} = -\frac{1}{\Omega^2}\left[1 + T_{\Omega}(0, \Omega) \frac{\Omega^{*2} - \Omega^2}{\gamma}\right]$$

Substitute both terms into the Taylor approximation and rearrange:

$$T(\hat{\mu}, \Omega) = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2} \left(1 + \mathcal{L}^{\tau}(\Omega)\right) \tag{A.25}$$

where $\mathcal{L}^{\tau}(\Omega) \equiv T_{\Omega}(0, \Omega) \frac{\Omega^{*2} - \Omega^2}{\gamma}$ measures the effect of uncertainty changes on the expected time and $\mathcal{L}_{\tau}(\Omega^*) = 0$.

3. **Approximation around** Ω^*. A first order Taylor approximation of $\mathcal{L}_\tau(\Omega)$ around Ω^* yields:

$$\mathcal{L}^\tau(\Omega) = \mathcal{L}^\tau(\Omega^*) + \mathcal{L}^\tau_\Omega(\Omega^*)(\Omega - \Omega^*) = -\frac{2\Omega^*}{\gamma} T_\Omega(0, \Omega^*)(\Omega - \Omega^*)$$

To characterize $T_\Omega(0, \Omega^*)$, take the partial derivative of (A.25) with respect to Ω and evaluate it at $(0, \Omega^*)$:

$$T_\Omega(\Omega^*, 0) = -\frac{2\bar{\mu}(\Omega)^2}{\Omega^{*3}}(1 - \mathcal{E}(\Omega^*))\Big(1 + \frac{2}{\gamma}\frac{\bar{\mu}(\Omega^*)^2}{\Omega^*}\Big)^{-1} = -\frac{2(1 - \mathcal{E}(\Omega^*))}{\Omega^{*2}}\Big(\frac{2\gamma(6\bar{\theta})^{1/2}}{\gamma + 2(6\bar{\theta})^{1/2}}\Big)$$

where $\mathcal{E}(\Omega^*)$ is the elasticity of the inaction region at Ω^*. Substitute back into $\mathcal{L}^\tau(\Omega)$ and arrive to

$$\mathcal{L}^\tau(\Omega) = 2\left(\frac{\Omega}{\Omega^*} - 1\right)(1 - \mathcal{E}(\Omega^*))\left(\frac{2\gamma(6\bar{\theta})^{1/2}}{\gamma + 2(6\bar{\theta})^{1/2}}\right)$$

Finally, we arrive at the result

$$T(\hat{\mu}, \Omega) = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}\left[1 + A\left(\frac{\Omega}{\Omega^*} - 1\right)\right]$$

where $A \equiv 2(1 - \mathcal{E}(\Omega^*))\left(\frac{2\gamma(6\bar{\theta})^{1/2}}{\gamma+2(6\bar{\theta})^{1/2}}\right)$ is a positive constant since the elasticity $\mathcal{E}(\Omega^*)$ is lower than unity. Furthermore, A is close to zero for small menu costs and large signal noise, as in our calibration.

4. **Decreasing and convex in uncertainty.** The expected time *between* price changes is equal to $T(0, \Omega)$:

$$T(0, \Omega) = \frac{\bar{\mu}(\Omega)^2}{\Omega^2}\left[1 + A\left(\frac{\Omega}{\Omega^*} - 1\right)\right]$$

Its first derivative with respect to uncertainty is given by:

$$\frac{\partial T(0, \Omega)}{\partial \Omega} = \frac{\bar{\mu}(\Omega)^2}{\Omega^3}\left(2(\mathcal{E}(\Omega) - 1)\left[1 + A\left(\frac{\Omega}{\Omega^*} - 1\right)\right] + A\frac{\Omega}{\Omega^*}\right)$$

If A is close to zero (as it is the case with small menu costs and large signal noise) we obtain:

$$\frac{\partial T(0, \Omega)}{\partial \Omega} = -2\frac{\bar{\mu}(\Omega)^2}{\Omega^3}(1 - \mathcal{E}(\Omega)) < 0$$

which is negative because the elasticity $\mathcal{E}(\Omega)$ is lower than unity. Finally, the second derivative

$$\frac{\partial^2 T(0, \Omega)}{\partial \Omega^2} = 4\frac{\bar{\mu}(\Omega)^2}{\Omega^4}\left[\left(\frac{3}{2} - \mathcal{E}(\Omega)\right)(1 - \mathcal{E}(\Omega)) + \frac{\Omega}{2}\mathcal{E}'(\Omega))\right] > 0$$

which is positive since the elasticity is lower than unity and increasing in uncertainty. Therefore, the expected time is decreasing and convex in uncertainty.

□

Proposition 6 (**Uncertainty and Frequency**). *The following relationship between uncertainty dispersion, average price duration, and price change dispersion holds:*

$$\mathbb{E}[\Omega^2] = \frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]} \tag{16}$$

Proof. See Proposition 1 in Álvarez, Le Bihan and Lippi (2014) for a derivation of this result for the case of fixed uncertainty $\Omega_t = \sigma$. Here we extend their proof for the case of stochastic uncertainty; most steps are analogous to theirs.

Recall that markup gap estimates follow $d\hat{\mu}_t = \Omega_t dB_t$. Using Itō's Lemma, we have that $d(\hat{\mu}_t^2) = \Omega_t^2 dt + 2\mu_t \Omega_t dB_t$. Therefore $d(\hat{\mu}_t^2) - \Omega_t^2 dt$ is a martingale. Let initial conditions be $(\mu_0, \Omega_0) = (0, \tilde{\Omega})$. Then using the Optional Stopping (or Doob's Sampling) Theorem, which says that the expected value of a martingale at a stopping time is equal to the expected value of its initial value (zero in our case), we have that

$$\begin{aligned}
\mathbb{E}\left[\hat{\mu}_\tau^2 - \int_0^\tau \Omega_s^2 ds \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] &= \mu_0^2 - \int_0^0 \Omega_s^2 ds = 0 \\
\mathbb{E}\left[\hat{\mu}_\tau^2 \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] &= \mathbb{E}\left[\int_0^\tau \Omega_s^2 ds \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right]
\end{aligned}$$

Now we will integrate both sides over different initial states. Since $\mu_0 = 0$ always at the stopping time, we only need to integrate over initial uncertainty $\tilde{\Omega}$ using the renewal density $r(\Omega)$, which is the distribution of uncertainty of adjusting firms.

- Integrating the LHS we obtain the unconditional (or cross-sectional) variance of price changes (recall that price changes are equal to markup gap estimates at adjustment, and that the mean price change is zero):

$$\int_0^\infty \mathbb{E}\left[\hat{\mu}_\tau^2 \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] r(\tilde{\Omega}) d\tilde{\Omega} = \mathbb{E}[(\Delta p)^2] = \mathbb{V}[(\Delta p)] \tag{A.26}$$

- Following Stokey (2009) the RHS is equal to the expected local time $\mathbb{L}$ for the payoff function Ω_t^2, which measures the expected amount of time the process has spent at state $(\hat{\mu}, \Omega)$ given initial condition $(0, \tilde{\Omega})$:

$$\mathbb{E}\left[\int_0^\tau \Omega_s^2 ds \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] = \int_{\hat{\mu}, \Omega} \mathbb{L}(\hat{\mu}, \Omega; 0, \tilde{\Omega}) \Omega^2 d\hat{\mu} d\Omega$$

  This allows us to express the expectation in state domain instead of the time domain. Again we integrate over all initial conditions using the renewal distribution:

$$\begin{aligned}
\int_0^\infty \left(\int_{\hat{\mu}, \Omega} \mathbb{L}(0, \tilde{\Omega}; \hat{\mu}, \Omega) \Omega^2 d\hat{\mu} d\Omega\right) r(\tilde{\Omega}) d\tilde{\Omega} &= \int_{\hat{\mu}, \Omega} \left(\int_0^\infty (\hat{\mu}, \Omega; 0, \tilde{\Omega})) r(\tilde{\Omega}) d\tilde{\Omega}\right) \Omega^2 d\hat{\mu} d\Omega \\
&= \mathbb{E}[\tau] \int_{\hat{\mu}, \Omega} \left(\int_0^\infty \frac{\mathbb{L}(\hat{\mu}, \Omega; 0, \tilde{\Omega})}{\mathbb{E}[\tau]} r(\tilde{\Omega}) d\tilde{\Omega}\right) \Omega^2 d\hat{\mu} d\Omega \\
&= \mathbb{E}[\tau] \int_{\hat{\mu}, \Omega} \Omega^2 f(\hat{\mu}, \Omega) d\hat{\mu} d\Omega \\
&= \mathbb{E}[\tau] \mathbb{E}[\Omega^2] \qquad \text{(A.27)}
\end{aligned}$$

  where in the first equality we have changed the order of integration, in the second equality we multiply and divide by the expected stopping time $\mathbb{E}[\tau]$ to normalize the local time, and in the third equality we use the relationship between the joint probability density $f(\hat{\mu}, \Omega)$ and the normalized local time $f(\hat{\mu}, \Omega) = \int_0^\infty \frac{\mathbb{L}(0, \tilde{\Omega}; \hat{\mu}, \Omega)}{\mathbb{E}[\tau]} r(\tilde{\Omega}) d\tilde{\Omega}$, and in the fourth equality we recognize the unconditional (or cross-sectional) second moment of Ω.

Putting together (A.26) and (A.27) we get the result:

$$\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]} = \mathbb{E}[\Omega^2]$$

□

Proposition 7 (**Conditional Hazard Rate).** *Without loss of generality, assume the last price change occurred at $t = 0$ and let $\Omega_0 > \sigma_f$ be the initial level of uncertainty. The inaction region is constant $\bar{\mu}(\Omega_\tau) = \bar{\mu}_0$ and there are no infrequent shocks ($\lambda = 0$). Denote derivatives with respect to τ with a prime ($h'_\tau \equiv \partial h/\partial\tau$).*

1. *The estimate's unconditional variance, denoted by $\mathcal{V}_\tau(\Omega_0)$, is given by:*

$$\mathcal{V}_\tau(\Omega_0) = \sigma_f^2\tau + \mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) \tag{17}$$

where $\mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) \equiv \gamma(\Omega_0 - \Omega_\tau)$, with $\mathcal{L}_0^{\mathcal{V}}(\Omega_0) = 0$, $\lim_{\tau\to\infty} \mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) = \gamma(\Omega_0 - \sigma_f)$, and it is equal to:

$$\mathcal{L}_\tau^{\mathcal{V}}(\Omega_0) = \gamma\Omega_0 - \gamma\sigma_f \left(\frac{\frac{\Omega_0}{\sigma_f} + \tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)} \right)$$

2. *$\mathcal{V}_\tau(\Omega_0)$ is increasing and concave in duration τ: $\mathcal{V}'_\tau(\Omega_0) > 0$ and $\mathcal{V}''_\tau(\Omega_0) < 0$. Furthermore, the following cross derivatives with initial uncertainty are positive:*

$$\frac{\partial \mathcal{V}_\tau(\Omega_0)}{\partial\Omega_0} > 0, \qquad \frac{\partial \mathcal{V}'_\tau(\Omega_0)}{\partial\Omega_0} > 0, \qquad \frac{\partial |\mathcal{V}''_\tau(\Omega_0)|}{\partial\Omega_0} > 0$$

3. *The hazard of adjusting the price at date τ, conditional on Ω_0, is characterized by:*

$$h_\tau(\Omega_0) = \frac{\pi^2}{8} \underbrace{\frac{\mathcal{V}'_\tau(\Omega_0)}{\bar{\mu}_0^2}}_{\text{decreasing in } \tau} \underbrace{\Psi\left(\frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)}_{\text{increasing in } \tau} \tag{18}$$

where $\Psi(x) \geq 0$, $\Psi(0) = 0$, $\Psi'(x) > 0$, $\lim_{x\to\infty}\Psi(x) = 1$, first convex then concave, and it is given by:

$$\Psi(x) = \frac{\sum_{j=0}^{\infty} \alpha_j \exp(-\beta_j x)}{\sum_{j=0}^{\infty} \frac{1}{\alpha_j}\exp(-\beta_j x)}, \quad \alpha_j \equiv (-1)^j(2j+1), \quad \beta_j \equiv \frac{\pi^2}{8}(2j+1)^2$$

4. *There exists a date $\tau^*(\Omega_0)$ such that the slope of the hazard rate is negative for $\tau > \tau^*(\Omega_0)$; and $\tau^*(\Omega_0)$ is decreasing in Ω_0.*

Proof. Assume $\lambda = 0$, initial conditions $(\hat{\mu}_0, \Omega) = (0, \Omega_0)$, and a constant inaction region at $\bar{\mu}_0 \equiv \bar{\mu}(\Omega) = \bar{\mu}(\Omega_0)$. Without loss of generality, we assume the last price change occurred at $t = 0$. First we derive expressions for two objects that will be part of the estimate's unconditional variance: the state's unconditional variance $\mathbb{E}_0[\mu_\tau^2]$ and the estimate's conditional variance Σ_τ. All moments are conditional on the initial conditions, but we do not make it explicit for simplicity.

1. **State's unconditional variance** Since the state evolves as $d\mu_\tau = \sigma_f dW_\tau$, we have that $\mu_\tau = \mu_0 + \sigma_f W_\tau$, with $W_0 = 0$ and $\mu_0 \sim \mathcal{N}(0, \Sigma_0)$. Therefore, the state's unconditional variance at time τ (after the last price change at $t = 0$) is given by:

$$\mathbb{E}_0[\mu_\tau^2] = \mathbb{E}_0[(\mu_0 + \sigma_f W_\tau)^2] = \mathbb{E}_0[\mu_0^2 + 2\mu_0\sigma_f\mathbb{E}_0[(W_\tau - W_0)] + \sigma_f^2\mathbb{E}_0[(W_\tau - W_0)^2] = \mathbb{E}_0[\mu_0^2] + \sigma_f^2\tau = \Sigma_0 + \sigma_f^2\tau \tag{A.28}$$

where we have use the properties of the Wiener process.

2. **Estimate's conditional variance**. The conditional forecast variance evolves as $d\Sigma_\tau = \left(\sigma_f^2 - \frac{\Sigma^2}{\gamma^2}\right)d\tau$. Assuming an initial condition Σ_0 such that $\Sigma_0 > \gamma\sigma_f$, the general solution to the differential equation is given by

$$\Sigma_\tau = \sigma_f\gamma\tanh\left[\sigma_f\gamma c + \frac{\sigma_f}{\gamma}\tau\right]$$

Evaluation at the initial condition, we get $\Sigma_0 = \sigma_f\gamma\tanh\left[\gamma\sigma_f c\right]$ and therefore $c = \frac{1}{\sigma_f\gamma}\tanh^{-1}\left(\frac{\Sigma_0}{\sigma_f\gamma}\right)$. Back into (2) and using properties of the hyperbolic tangent,

$$\Sigma_\tau = \sigma_f\gamma\tanh\left[\tanh^{-1}\left(\frac{\Sigma_0}{\sigma_f\gamma}\right) + \frac{\sigma_f}{\gamma}\tau\right] = \sigma_f\gamma\left(\frac{\frac{\Sigma_0}{\sigma_f\gamma} + \tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Sigma_0}{\sigma_f\gamma}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}\right) \tag{A.29}$$

Since $\tanh(0) = 0$ and $\tanh(+\infty) = 1$ we confirm that $\Sigma_\tau = \Sigma_0$ at $\tau = 0$ and $\lim_{\tau\to\infty}\Sigma_\tau = \sigma_f\gamma$.

3. **Estimate's unconditional variance.** Recall that the estimate follows $d\hat{\mu}_\tau = \Omega_\tau d\hat{Z}_\tau$. Since $\lambda = 0$, uncertainty evolves deterministically as $d\Omega_\tau = \frac{1}{\gamma}(\sigma_f^2 - \Omega_\tau^2)$. Given the initial condition $\hat{\mu}_0 = 0$, the solution to the forecast equation is $\hat{\mu}_\tau = \int_0^\tau \Omega_s d\hat{Z}_s$. By definition of Itō's integral $\int_0^\tau \Omega_s d\hat{Z}_s = \lim_{(\tau_{i+1}-\tau_i)\to 0}\sum_{\tau_i}\Omega_{\tau_i}(\hat{Z}_{\tau_{i+1}} - \hat{Z}_{\tau_i})$. The increments' Normality and the fact that Ω_{τ_i} is deterministic imply that for each τ_i, $\Omega_{\tau_i}(\hat{Z}_{\tau_{i+1}} - \hat{Z}_{t_i})$ is Normally distributed as well. Since the limit of Normal variables is Normal, we have that markup gap's estimate at date τ, given information set $\mathcal{I}_0$, is also Normally distributed. Let $\mathcal{V}_\tau \equiv \mathbb{E}_0[\hat{\mu}_\tau^2]$ denote the estimate's unconditional variance, then $\hat{\mu}_\tau|\mathcal{I}_0 \sim \mathcal{N}(0, \mathcal{V}_\tau)$. To characterize $\mathcal{V}_\tau$, start from its definition and add and subtract μ_t:

$$\mathcal{V}_\tau \equiv \mathbb{E}_0[\hat{\mu}_\tau^2] = \mathbb{E}_0[\mu_\tau^2] + \mathbb{E}_0[(\hat{\mu}_\tau - \mu_t)^2] - 2\mathbb{E}_0[(\hat{\mu}_\tau - \mu_\tau)\mu_\tau] = \mathbb{E}_0[\mu_\tau^2] - \Sigma_\tau \tag{A.30}$$

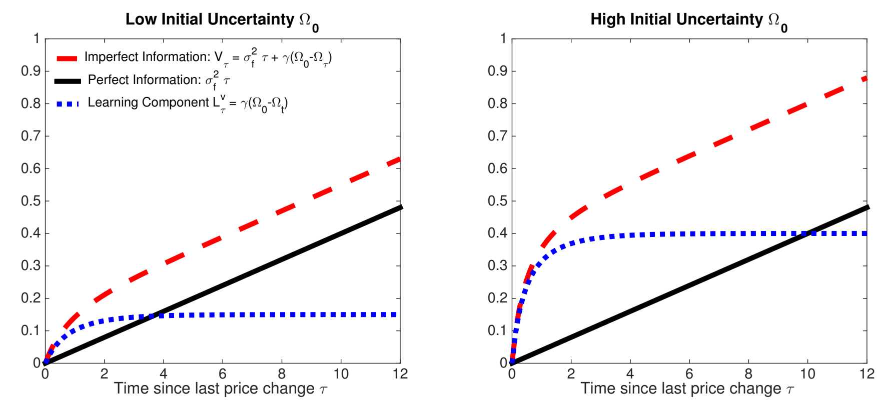
where we that $\mathbb{E}_0[(\hat{\mu}_\tau - \mu_\tau)\mu_\tau] = \mathbb{E}_0[(\hat{\mu}_\tau - \mu_\tau)^2] = \Sigma_t$, implied by the orthogonality of the innovation and the forecast: $\mu_\tau - \hat{\mu}_\tau \perp \hat{\mu}_\tau$. Substituting expressions (A.28) and (A.29) into (A.30) and using $\Omega_\tau = \gamma\Sigma_\tau$, we get:

$$\mathcal{V}_\tau = \sigma_f^2\tau + \gamma\left(\Omega_0 - \Omega_\tau\right) = \sigma_f^2\tau + \gamma\left(\Omega_0 - \sigma_f\left(\frac{\frac{\Omega_0}{\sigma_f} + \tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}\right)\right) = \sigma_f^2\tau + \mathcal{L}_\tau^{\mathcal{V}} \tag{A.31}$$

where we define the learning component as:

$$\mathcal{L}_\tau^{\mathcal{V}} \equiv \gamma\left(\Omega_0 - \sigma_f\left(\frac{\frac{\Omega_0}{\sigma_f} + \tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}\right)\right)$$

The hyperbolic tangent function is defined as $\tanh(x) \equiv \frac{e^x - e^{-x}}{e^x + e^{-x}}$, and has the following properties: $\tanh(0) = 0$, $\lim_{x\to\pm\infty}\tanh(x) = \pm 1$, $\tanh'(x) = 1 - \tanh^2(x)$.



Useful derivatives The first and second derivatives of the learning component with respect to τ are given by:

$$\begin{aligned}
\frac{\partial\mathcal{L}_\tau^{\mathcal{V}}}{\partial\tau} &= \sigma_f^2\left(\frac{\Omega_0}{\sigma_f} - 1\right)\frac{1 - \tanh^2\left(\frac{\sigma_f}{\gamma}\tau\right)}{\left[1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)\right]^2} > 0 \\
\frac{\partial^2\mathcal{L}_\tau^{\mathcal{V}}}{\partial\tau^2} &= -\frac{2\sigma_f}{\gamma}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)\left[1 + \frac{\Omega_0}{\sigma_f}\frac{1 - \tanh^2\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)}\right]\frac{\partial\mathcal{L}_\tau^{\mathcal{V}}}{\partial\tau} < 0
\end{aligned}$$

The derivative of the learning component with respect to uncertainty is:

$$\frac{\partial\mathcal{L}_\tau^{\mathcal{V}}}{\partial\Omega_0} = \gamma - \frac{1 - \tanh^2\left(\frac{\sigma_f}{\gamma}\tau\right)}{\left[1 + \frac{\Omega_0}{\sigma_f}\tanh\left(\frac{\sigma_f}{\gamma}\tau\right)\right]^2}, \quad \text{positive for large } \gamma, \text{ large } \Omega_0, \text{ and large } \tau$$

Furthermore, the following relationship and signs hold:

$$\frac{\partial\mathcal{L}_\tau^{\mathcal{V}}}{\partial\tau} = \sigma_f^2\left(\frac{\Omega_0}{\sigma_f} - 1\right)\left(\gamma - \frac{\partial\mathcal{L}_\tau^{\mathcal{V}}}{\partial\Omega_0}\right), \qquad \frac{\partial^2\mathcal{L}_\tau^{\mathcal{V}}}{\partial\tau\partial\Omega_0} > 0, \qquad \left|\frac{\partial^3\mathcal{L}_\tau^{\mathcal{V}}}{\partial\tau^2\partial\Omega_0}\right| > 0$$

4. **Stopping time distribution.** Let $F(\sigma_f^2\tau, \bar{\mu}_0)$ be the cumulative distribution of stopping times obtained from a problem with perfect information which considers a Brownian motion with unconditional variance of $\sigma_f^2\tau$, initial condition 0, and a symmetric inaction region $[-\bar{\mu}_0, \bar{\mu}_0]$. Following Kolkiewicz (2002) and Álvarez, Lippi and Paciello (2011)'s Online Appendix, the density of stopping times is given by:

$$f(\tau) = \frac{\pi}{2}x'(\tau)\sum_{j=0}^{\infty}\alpha_j\exp\left(-\beta_j x(\tau)\right), \qquad \text{where} \quad x(\tau) \equiv \frac{\sigma_f^2\tau}{\bar{\mu}_0^2}, \quad \alpha_j \equiv (2j+1)(-1)^j, \quad \beta_j \equiv (2j+1)^2\frac{\pi^2}{8}$$

The process $x(\tau)$ is equal to the ratio of volatility and the width of the inaction region. Since we assumed constant inaction regions, x only changes with volatility. In our case, the estimate's unconditional variance is given by $\mathcal{V}_\tau(\Omega_0)$. Using a change of variable, the distribution of stopping times becomes $F\left(\mathcal{V}_\tau(\Omega_0), \bar{\mu}_0\right)$ with density $f(\tau|\Omega_0) = f(\mathcal{V}_\tau(\Omega_0), \bar{\mu}_0)$. We can apply the previous formula using $x \equiv \frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}$ and the same sequences of α_j and β_j.

5. **Hazard rate.** Given the stopping time distribution, the conditional hazard rate is computed using its definition:

$$h_\tau(\Omega_0) \equiv \frac{f(\tau|\Omega_0)}{\int_\tau^\infty f(s|\Omega)ds} = \frac{f(\mathcal{V}_\tau(\Omega_0), \bar{\mu}_0)}{\int_\tau^\infty f(\mathcal{V}_s(\Omega_0), \bar{\mu}_0)ds} = \frac{\mathcal{V}'_\tau(\Omega_0)\sum_{j=0}^\infty \alpha_j \exp\left(-\beta_j \frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)}{\int_\tau^\infty \mathcal{V}'_s(\Omega_0)\sum_{j=0}^\infty \alpha_j \exp\left(-\beta_j \frac{\mathcal{V}_s(\Omega_0)}{\bar{\mu}_0^2}\right)ds} \tag{A.32}$$

Let $u_j(s) \equiv \alpha_j \exp\left(-\beta_j \frac{\mathcal{V}_s(\Omega_0)}{\bar{\mu}_0^2}\right)$, then $du_j(s) \equiv \frac{-\alpha_j\beta_j}{\bar{\mu}_0^2}\mathcal{V}'_s(\Omega_0)\exp\left(-\beta_j \frac{\mathcal{V}_s(\Omega_0)}{\bar{\mu}_0^2}\right)ds$. Exchanging the summation with the integral, the denominator is equal to:

$$\sum_{j=0}^\infty \frac{-\bar{\mu}_0^2}{\beta_j}\int_\tau^\infty du_j(s)ds = \sum_{j=0}^\infty \frac{-\bar{\mu}_0^2}{\beta_j} u_j(s)\Big|_\tau^\infty = \sum_{j=0}^\infty \frac{\bar{\mu}_0^2}{\beta_j}u_j(\tau) = \bar{\mu}_0^2 \sum_{j=0}^\infty \frac{\alpha_j}{\beta_j}\exp\left(-\beta_j\frac{\mathcal{V}_s(\Omega_0)}{\bar{\mu}_0^2}\right) = \frac{8\bar{\mu}_0^2}{\pi^2}\sum_{j=0}^\infty \frac{1}{\alpha_j}\exp\left(-\beta_j\frac{\mathcal{V}_s(\Omega_0)}{\bar{\mu}_0^2}\right)$$

where in the last equality we use $\frac{\alpha_j}{\beta_j} = \frac{(2j+1)(-1)^j}{(2j+1)^2\frac{\pi^2}{8}} = \frac{8}{\pi^2}(2j+1)^{-1}(-1)^j = \frac{8}{\pi^2}\frac{1}{\alpha_j}$. Substituting back into (A.32):

$$h_\tau(\Omega_0) = \frac{\pi^2}{8\bar{\mu}_0^2}\Psi\left(\frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)\mathcal{V}'_\tau(\Omega_0) \tag{A.33}$$

where we define $\Psi(x) \equiv \frac{\sum_{j=0}^\infty \alpha_j \exp(-\beta_j x)}{\sum_{j=0}^\infty \frac{1}{\alpha_j}\exp(-\beta_j x)}$ as in Álvarez, Lippi and Paciello (2011)'s Online Appendix. The function $\Psi(x)$ is increasing, first convex then concave, with $\Psi(0) = 0$ and $\lim_{x\to\infty}\Psi(x) = 1$.

6. **Hazard rate's slope.** Taking derivative of the hazard rate with respect to duration τ yields:

$$h'_\tau \propto \underbrace{\frac{\partial^2 \mathcal{L}_\tau^\mathcal{V}}{\partial\tau^2}}_{<0}\underbrace{\Psi\left(\frac{\mathcal{V}_\tau}{\bar{\mu}_0^2}\right)}_{\to 1} + \underbrace{\left(\frac{\sigma_f^2 + \frac{\partial\mathcal{L}_\tau^\mathcal{V}}{\partial\tau}}{\bar{\mu}_0^2}\right)^2}_{>0}\underbrace{\Psi'\left(\frac{\mathcal{V}_\tau}{\bar{\mu}_0^2}\right)}_{\to 0}$$

For small τ, Ψ's derivative is very large and the second positive term dominates; as τ increases, the function Ψ and its derivative Ψ' converge to 1 and 0 respectively, and therefore the first term – which is negative – dominates. By the Intermediate Value Theorem, there exists a $\tau^*(\Omega_0)$ such that the slope is zero.

Taking the cross-derivative with respect to uncertainty and using the equivalence between derivatives stated above:

$$\frac{\partial h'_\tau}{\partial\Omega_0} \propto \underbrace{\Psi\left(\frac{\mathcal{V}_\tau}{\bar{\mu}_0^2}\right)}_{\to 1}\underbrace{\frac{\partial^3\mathcal{L}_\tau^\mathcal{V}}{\partial\tau^2\partial\Omega_0}}_{<0} + \underbrace{\Psi''\left(\frac{\mathcal{V}_\tau}{\bar{\mu}_0^2}\right)}_{\to 0^-}\underbrace{\left(\frac{\sigma_f^2 + \frac{\partial\mathcal{L}_\tau^\mathcal{V}}{\partial\tau}}{\bar{\mu}_0^2}\right)^2\frac{\partial\mathcal{L}_\tau^\mathcal{V}}{\partial\Omega_0}\frac{1}{\bar{\mu}_0^2}}_{>0} + \underbrace{\Psi'\left(\frac{\mathcal{V}_\tau}{\bar{\mu}_0^2}\right)}_{\to 0}\frac{1}{\bar{\mu}_0^2}\left[\underbrace{\frac{\partial^2\mathcal{L}_\tau^\mathcal{V}}{\partial\tau^2}\frac{\partial\mathcal{L}_\tau^\mathcal{V}}{\partial\Omega_0}}_{<0} + \underbrace{\frac{2\sigma_f^2}{\bar{\mu}_0^2}\frac{\partial^2\mathcal{L}_\tau^\mathcal{V}}{\partial\tau\partial\Omega_0}\left(1+\left(\frac{\Omega_0}{\sigma_f}-1\right)\left(\gamma - \frac{\partial\mathcal{L}_\tau^\mathcal{V}}{\partial\Omega_0}\right)\right)}_{>0}\right]$$

Since Ψ' and Ψ'' converge to 0 as τ increases, the first term dominates. Then the slope of the hazard rate becomes more negative as initial uncertainty increases. This means that the cutoff duration $\tau^*(\Omega_0)$ is decreasing with Ω_0.

□

Proposition 8 (**Renewal distribution**). *Let $f(\hat{\mu}, \Omega)$ be the joint density of markup gaps and uncertainty in the population of firms. Let $r(\Omega)$ be denote the density of uncertainty conditional on adjusting, or renewal density. Assume the inaction region is increasing in uncertainty (i.e. $\bar{\mu}'(\Omega) > 0$). Then we have the following results:*

1. *For each $(\hat{\mu}, \Omega)$, we can write the joint density as $f(\hat{\mu}, \Omega) = h(\Omega)g(\hat{\mu}, \Omega)$, where $g(\hat{\mu}, \Omega)$ is the density of markup gap estimates conditional on uncertainty and $h(\Omega)$ is the marginal density of uncertainty.*
2. *The ratio between the renewal and marginal densities of uncertainty is approximated by*

$$\frac{r(\Omega)}{h(\Omega)} \propto |g_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega)|\Omega^2 \tag{19}$$

where $g(\mu, \Omega)$ solves the following differential equation $\frac{\Omega^2 - \Omega^{*2}}{\gamma} g_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} g_{\hat{\mu}^2}(\hat{\mu}, \Omega) = 0$ with border conditions: $g(\bar{\mu}(\Omega), \Omega) = 0$ and $\int_{-\bar{\mu}(\Omega)}^{\bar{\mu}(\Omega)} g(\mu, \Omega) d\mu = 1$.

3. *If $\Omega = \Omega^*$, then the ratio is proportional to the inverse of the expected time between price adjustments. Then if the inaction region's elasticity to uncertainty is lower than unity, the ratio is an increasing function of uncertainty:*

$$\frac{r(\Omega^*)}{h(\Omega^*)} \propto \frac{\Omega^{*2}}{2\bar{\mu}(\Omega^*)^2} = \frac{1}{2\mathbb{E}[\tau|(0, \Omega^*)]} \tag{20}$$

Proof. The strategy for the proof is as follows. We derive the Kolmogorov Forward Equation (KFE) of the joint ergodic distribution using the adjoint operator. Then we find the zeros of the KFE to characterize the ergodic distribution.

1. **Joint distribution.** Let $f(\hat{\mu}, \Omega) : [-\infty, \infty] \times [\sigma_f, \infty] \to \mathbb{R}$ be the ergodic density of markup estimates and uncertainty. Define the region:

$$\mathcal{R}^\circ \equiv \left\{(\hat{\mu}, \Omega) \in [-\infty, \infty] \times [\sigma_f, \infty] \quad \text{such that} \quad |\hat{\mu}| < \bar{\mu}(\Omega) \ \& \ \hat{\mu} \neq 0\right\} \tag{A.34}$$

where $\bar{\mu}(\Omega)$ is the border of the inaction region. Thus $\mathcal{R}^\circ$ is equal to the continuation region except $\hat{\mu} \neq 0$. Then the function f has the following properties:

a) f is continuous

b) f is zero outside the continuation region. Given Ω, $f(x, \Omega) = 0 \ \forall x \notin (-\bar{\mu}(\Omega), \bar{\mu}(\Omega))$. In particular, it is zero at the borders of the inaction region:

$$f(-\bar{\mu}(\Omega), \Omega) = 0 = f(\bar{\mu}(\Omega), \Omega), \quad \forall \Omega$$

c) f is a density: $\forall (\hat{\mu}, \Omega) \in \mathcal{R}^\circ$, we have that $f(\hat{\mu}, \Omega) \geq 0$ and $\int_{\Omega \geq \sigma_f} \int_{|\hat{\mu}| \leq \bar{\mu}(\Omega)} f(\hat{\mu}, \Omega) d\hat{\mu} d\Omega = 1$

d) For any state $(\hat{\mu}, \Omega) \in \mathcal{R}^\circ$, f is a zero of the Kolmogorov Forward Equation (KFE):

$$A^* f(\hat{\mu}, \Omega) = 0$$

Substituting the adjoint operator A^* obtained in (A.2) we write the KFE as:

$$-\frac{\sigma_f^2 - \Omega^2}{\gamma} f_\Omega(\hat{\mu}, \Omega) + \frac{2\Omega}{\gamma} f(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda \left[f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu}, \Omega)\right] = 0 \tag{A.35}$$

We compute $f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right)$ with a first order Taylor approximation: $f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) \approx f(\hat{\mu}, \Omega) - \frac{\sigma_u^2}{\gamma} f_\Omega(\hat{\mu}, \Omega)$. Substituting this approximation, collecting terms, and using the definition of fundamental uncertainty Ω^*, the KFE becomes:

$$\frac{2\Omega}{\gamma} f(\hat{\mu}, \Omega) + \frac{\Omega^2 - \Omega^{*2}}{\gamma} f_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu}, \Omega) = 0 \tag{A.36}$$

with two border conditions:

$$\forall \Omega \quad f(|\bar{\mu}(\Omega)|, \Omega) = 0 \quad ; \quad \int_{\Omega \geq \sigma_f} \int_{|\hat{\mu}| \leq \bar{\mu}(\Omega)} f(\hat{\mu}, \Omega) d\hat{\mu} d\Omega = 1 \tag{A.37}$$

2. **Marginal density of uncertainty** Let $h(\Omega) : [\sigma_f, \infty] \to \mathbb{R}$ be the uncertainty's ergodic density; it solves the following KFE

$$A^* h = \frac{2\Omega}{\gamma} h(\Omega) + \frac{\Omega^2 - \Omega^{*2}}{\gamma} h_\Omega(\Omega) = 0$$

and a border condition $\lim_{\Omega \to \infty} h(\Omega) = 0$.

3. **Factorization of f.** For each $(\hat{\mu}, \Omega)$, guess that we can write f as a product of the ergodic density of uncertainty h and a function g as follows:

$$f(\hat{\mu}, \Omega) = h(\Omega) g(\hat{\mu}, \Omega) \tag{A.38}$$

Substituting (A.38) into (A.36) and rearranging

$$
\begin{aligned}
0 &= \frac{2\Omega}{\gamma}h(\Omega)g(\hat{\mu},\Omega) + \frac{\Omega^2-\Omega^{*2}}{\gamma}\left[h_\Omega(\Omega)g(\hat{\mu},\Omega)+h(\Omega)g_\Omega(\hat{\mu},\Omega)\right] + \frac{\Omega^2}{2}h(\Omega)g_{\hat{\mu}^2}(\hat{\mu},\Omega) \\
&= g(\hat{\mu},\Omega)\underbrace{\left[\frac{2\Omega}{\gamma}h(\Omega)+\frac{\Omega^2-\Omega^{*2}}{\gamma}h_\Omega(\Omega)\right]}_{\text{KFE for } h} + h(\Omega)\left[\frac{\Omega^2-\Omega^{*2}}{\gamma}g_\Omega(\hat{\mu},\Omega)+\frac{\Omega^2}{2}h(\Omega)g_{\hat{\mu}^2}(\hat{\mu},\Omega)\right] \\
&= \frac{\Omega^2-\Omega^{*2}}{\gamma}g_\Omega(\hat{\mu},\Omega)+\frac{\Omega^2}{2}g_{\hat{\mu}^2}(\hat{\mu},\Omega)
\end{aligned}
$$

where in the second line we regroup terms and recognize the KFE for h, in the third line we set the KFE of h equal to zero because it is uncertainty's ergodic density and divide by h as it is assumed to be positive. To obtain the border conditions for g, substitute the decomposition (A.38) into (A.37):

$$
\forall\Omega \quad h(\Omega)g(|\bar{\mu}(\Omega)|,\Omega)=0 \quad ; \quad \int_{\Omega\geq\sigma_f}\int_{|\hat{\mu}|\leq\bar{\mu}(\Omega)} h(\Omega)g(\hat{\mu},\Omega)d\hat{\mu}d\Omega = 1 \tag{A.39}
$$

Since $h>0$, we can eliminate it in the first condition and get a border condition for g:

$$
g(|\bar{\mu}(\Omega)|,\Omega)=0
$$

Then assume that for each Ω, g integrates to one. Use this assumption into the second condition:

$$
\int_{\Omega\geq\sigma_f} h(\Omega)\left[\int_{|\hat{\mu}|\leq\bar{\mu}(\Omega)} g(\hat{\mu},\Omega)d\hat{\mu}\right]d\Omega = \int_{\Omega\geq\sigma_f} h(\Omega)d\Omega = 1
$$

Therefore, by the factorization method, the ergodic distribution h is also the marginal density $h(\Omega)=\int_{|\hat{\mu}|\leq\bar{\mu}(\Omega)} f(\hat{\mu},\Omega)d\hat{\mu}$ and g is the density of markup gap estimates conditional on uncertainty $g(\hat{\mu},\Omega)=f(\hat{\mu}|\Omega)=\frac{f(\hat{\mu},\Omega)}{h(\Omega)}$.

4. **Renewal density** The renewal density is the distribution of firm uncertainty conditional on a price adjustment. For each unit of time, the fraction of firms that adjusts at given uncertainty level is given by three terms (the terms multiplied by 2 take into account the symmetry of the distribution around a zero markup gap):

$$
r(\Omega) \propto 2f(\bar{\mu}(\Omega),\Omega)\frac{\sigma_f^2-\Omega^2}{\gamma} + \lambda\int_{-\bar{\mu}(\Omega-\sigma_u^2/\gamma)}^{\bar{\mu}(\Omega-\sigma_u^2/\gamma)} f\left(\mu,\Omega-\frac{\sigma_u^2}{\gamma}\right) I\left(\hat{\mu}>\bar{\mu}\left(\Omega\right)\right)d\mu d\Omega + 2|f_{\hat{\mu}}(\bar{\mu}(\Omega),\Omega)|\frac{\Omega^2}{2} \tag{A.40}
$$

The first term counts price changes of firms at the border of the inaction region that suffer a deterministic decrease in uncertainty; by the border condition $f(\bar{\mu}(\Omega),\Omega)=0$, this term is equal to zero. The second term counts price changes due to jumps in uncertainty. These firms had an uncertainty level of $\Omega-\frac{\sigma_u^2}{\gamma}$ right before the jump; under the assumption that $\bar{\mu}(\Omega)$ is increasing in uncertainty, this term is also equal to zero since all markup estimates that were inside the initial inaction region remain inside the new inaction region. The last term counts price changes of firms at the border of the inaction region that suffer either a positive or negative change in the markup gap estimate (hence the absolute value). This term is the only one different from zero. Substituting the factorization of f, we obtain a simplified expression for the renewal distribution in terms of g:

$$
\frac{r(\Omega)}{h(\Omega)} \propto |g_{\hat{\mu}}(\bar{\mu}(\Omega),\Omega)|\Omega^2 \tag{A.41}
$$

5. **Characterize g when $\Omega=\Omega^*$.** If $\Omega=\Omega^*$, the markup gap conditional distribution g can be further characterized:

$$
g_{\mu^2}(\hat{\mu},\Omega^*)=0; \qquad g(\bar{\mu}(\Omega^*),\Omega^*)=0; \qquad \int_{-\bar{\mu}(\Omega^*)}^{\bar{\mu}(\Omega^*)} g(\hat{\mu},\Omega^*)d\hat{\mu}=1 \qquad g\in\mathbb{C} \tag{A.42}
$$

To solve this equation, integrate twice with respect to $\hat{\mu}$: $g(\hat{\mu},\Omega^*)=|C|\hat{\mu}+|D|$. To determine the constants $|C|$ and $|D|$, we use the border conditions:

$$
\begin{aligned}
0 &= g(\bar{\mu}(\Omega^*),\Omega^*) = |C|\bar{\mu}(\Omega^*)+|D| \\
1 &= \int_{-\bar{\mu}(\Omega^*)}^{\bar{\mu}(\Omega^*)} g(\hat{\mu},\Omega^*)d\mu = \int_{-\bar{\mu}(\Omega^*)}^{\bar{\mu}(\Omega^*)} (|C|\hat{\mu}+|D|)d\mu = \left(\frac{|C|}{2}\hat{\mu}^2+|D|\hat{\mu}\right)\Big|_{-\bar{\mu}(\Omega^*)}^{\bar{\mu}(\Omega^*)} = 2\bar{\mu}(\Omega^*)|D|
\end{aligned}
$$

From the second equality, we get that

$$
D=\frac{1}{2\bar{\mu}(\Omega^*)}
$$

Then substituting in the first equality:

$$
|C| = -\frac{|D|}{\bar{\mu}(\Omega^*)} = -\frac{1}{2\bar{\mu}(\Omega^*)^2}
$$

Lastly, since $g_{\mu^2}(\hat{\mu}, \Omega^*) \geq 0$, we obtain :

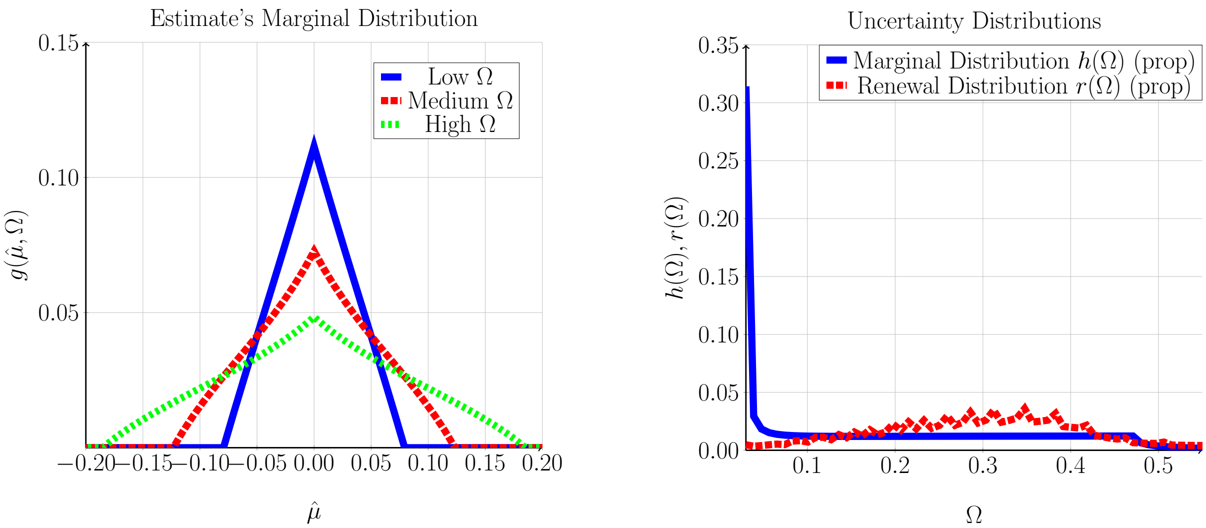
$$g(\mu, \hat{\Omega}) = \begin{cases} \frac{1}{2\bar{\mu}(\Omega^*)}\left(1 + \frac{\hat{\mu}}{\bar{\mu}(\Omega^*)}\right) & \text{if } \hat{\mu} \in [-\bar{\mu}(\hat{\Omega}), 0] \\ \frac{1}{2\bar{\mu}(\Omega^*)}\left(1 - \frac{\hat{\mu}}{\bar{\mu}(\Omega^*)}\right) & \text{if } \hat{\mu} \in (0, \bar{\mu}(\hat{\Omega})] \end{cases} \tag{A.43}$$

This is a triangular distribution in the $\hat{\mu}$ domain for each Ω (see next figure).

6. **Ratio when $\Omega = \Omega^*$.** By the previous result, the ratio of the renewal to marginal distributions at Ω^* is equal to:

$$\frac{r(\Omega^*)}{h(\Omega^*)} = \left|g_{\hat{\mu}}(\bar{\mu}(\Omega^*), \Omega^*)\right| \Omega^{*2} = \frac{\Omega^{*2}}{2\bar{\mu}(\Omega^*)^2} = \frac{1}{2\mathbb{E}[\tau|(0, \Omega^*)]} \tag{A.44}$$

Since the inaction region's elasticity to uncertainty is lower than unity, this ratio is increasing in uncertainty.



□

Proposition 9 (**Output Effects from Monetary and Uncertainty Shocks**). *Assume the economy is in steady state and it is hit with one-time unanticipated monetary shock of size δ, and firms only observe a fraction $\alpha \in [0,1]$ of it. Simultaneously, idiosyncratic firm uncertainty increases by $\kappa\bar{\Omega}$. Before the monetary and uncertainty aggregate shocks, firms' states are denoted by $(\hat{\mu}_{-1}, \Omega_{-1})$ distributed according to F.*

1. *Immediately after aggregate shocks arrive, but before idiosyncratic shocks do, markup estimates and uncertainty jump to $\hat{\mu}_0 = \hat{\mu}_{-1} - \alpha\delta$ and $\Omega_0 = \Omega_{-1} + \kappa\bar{\Omega}$. Before idiosyncratic shocks hit, forecast errors are random, and conditional on uncertainty, they are Normally distributed: $\varphi_0 \sim \mathcal{N}(-(1-\alpha)\delta, \gamma\Omega_0)$.*

2. *Let w be the future stream of pricing mistakes for a firm with state $(\hat{\mu}, \Omega, \varphi)$; it is computed recursively as*

$$w(\hat{\mu}, \Omega, \varphi) \quad = \quad \mathbb{E}\left[\int_0^\tau (\hat{\mu}_t + \varphi_t)\ dt + w(0, \Omega_\tau, \varphi_\tau) \Big| (\hat{\mu}_0, \Omega_0, \varphi_0) = (\hat{\mu}, \Omega, \varphi)\right] \tag{35}$$

subject to the following stochastic process:

$$d\hat{\mu}_t = \Omega_t \frac{\varphi_t}{\gamma} dt + \Omega_t dZ_t; \quad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t; \quad d\varphi_t = -\Omega_t \frac{\varphi_t}{\gamma} dt + \sigma_f dW_t + \sigma_u u_t dQ_t - \Omega_t dZ_t$$

3. *The total output response averages across all firms streams of pricing mistakes, taking into account the steady state distribution and the distribution of forecast errors:*

$$\mathcal{M}(\delta, \alpha, \kappa\bar{\Omega}) = -\int_{\hat{\mu},\Omega} \left[\int_{\varphi_0} w(\hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, x)\phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma(\Omega_{-1} + \kappa\bar{\Omega})}\right) d\varphi_0\right] dF(\hat{\mu}_{-1}, \Omega_{-1}) \tag{36}$$

4. *If $\alpha = 1$ (fully disclosed), then*

$$\mathcal{M}(\delta, 1, \kappa\bar{\Omega}) = -\int_{\hat{\mu},\Omega} \mathbb{E}\left[\int_0^\tau \hat{\mu}_t\ dt \Big| (\hat{\mu}_0, \Omega_0) = (\hat{\mu}_{-1} - \delta, \Omega_{-1} + \kappa\bar{\Omega})\right] dF(\hat{\mu}_{-1}, \Omega_{-1}) \tag{37}$$

subject to: $\quad d\hat{\mu}_t = \Omega_t d\hat{Z}_t; \quad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t.$

Proof. We divide the proof in five steps.

1. **Initial conditions**: A positive monetary shock of size δ translates as a downward jump in markups $\mu_0 = \mu_{-1} - \delta$. If the firms only incorporate a fraction α of the shock, then we have that markup estimates are adjusted by $\hat{\mu}_0 = \hat{\mu}_{-1} - \alpha\delta$. From Proposition 1 we have that, in the absence of the monetary shock, forecast errors are distributed Normally as $\varphi_t \sim N(0, \gamma\Omega_t)$. Therefore, at $t = 0$ before the idiosyncratic shocks are realized, we adjust the mean to take into account the knowledge about the monetary shock; the variance is not adjusted as all firms are affected in the same way: $\varphi_0 \sim \mathcal{N}(-(1-\alpha)\delta, \gamma\Omega_0)$. Finally, uncertainty gets amplified by a factor κ, thus $\Omega_0 = \Omega_{-1} + \kappa\bar{\Omega}$. Next we derive the law of motion for $(\hat{\mu}_t, \Omega_t, \varphi_t)$ for $t > 0$.

2. **State's stochastic process**: From equation (A.7), together with the definition of forecast errors $\varphi_t \equiv \mu_t - \hat{\mu}_t$, we can write the process for markup gap estimates in terms of forecast errors instead of the innovations' representation:

$$d\hat{\mu}_t = \Omega_t \left(\frac{\mu_t - \hat{\mu}_t}{\gamma} dt + dZ_t\right) = \Omega_t \frac{\varphi_t}{\gamma} dt + \Omega_t dZ_t$$

For forecast errors, we apply its definition and obtain: $d\varphi_t = \sigma_f dW_t + \sigma_t u_t dQ_t - \Omega_t \frac{\varphi_t}{\gamma} dt - \Omega_t dZ_t$. The process for uncertainty is the same as (A.6).

3. **Steady state and transition distribution:** Let F denote the firms' steady state distribution and let G be the initial cross-sectional distribution after the aggregate shock but before repricing. Apply Bayes' law:

$$G(\hat{\mu}_0, \Omega_0, \varphi_0) = G(\varphi_0 | \hat{\mu}_0, \Omega_0)\ G(\hat{\mu}_0, \Omega_0)$$

Given that $\varphi_0 \sim \mathcal{N}(-(1-\alpha)\delta, \gamma\Omega_0)$, we have that the conditional density of initial forecast errors is

$$G(\varphi_0 | \hat{\mu}_0, \Omega_0) = \Phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma\Omega_0}\right)$$

with $\Phi(\cdot)$ the distribution of a standard Normal. Then, since all markup gap estimates get shifted to the left and uncertainty to the right, the initial density for these states is a transformation of the steady state distribution:

$$G(\hat{\mu}_0, \Omega_0) = F\left(\hat{\mu}_{-1} + \alpha\delta, \Omega_{-1} - \kappa\bar{\Omega}\right)$$

Summarizing, the initial density after the aggregate shocks is given by:

$$G(\hat{\mu}_0, \Omega_0, \varphi_0) = \Phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma\Omega_0}\right) F\left(\hat{\mu}_{-1} + \alpha\delta, \Omega_{-1} - \kappa\bar{\Omega}\right)$$

4. **Recursive pricing mistakes**

Let τ_i the time of the i-th price change of firm with current state $(\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i})$ and define the function w as:

$$w(\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}) \equiv \mathbb{E}\left[\int_{\tau_i}^{\infty} (\hat{\mu}_t + \varphi_t) dt \,\middle|\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}\right] \tag{A.45}$$

subject to the stochastic process for the state. This function measures the stream of future pricing mistakes by the firm, which will produce output deviations from a frictionless case. Note that we can write w recursively:

$$\begin{aligned}
w(\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}) &= \mathbb{E}\left[\int_{\tau_i}^{\infty} (\hat{\mu}_t + \varphi_t) dt \,\middle|\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}\right] \\
&= \mathbb{E}\left[\int_{\tau_i}^{\tau_{i+1}} (\hat{\mu}_t + \varphi_t) dt + \mathbb{E}\left[\int_{\tau_{i+1}}^{\infty} (\hat{\mu}_t + \varphi_t) dt \,\middle| 0, \Omega_{\tau_{i+1}}, \varphi_{\tau_{i+1}}\right] \middle| \hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}\right] \\
&= \mathbb{E}\left[\int_{\tau_i}^{\tau_{i+1}} (\hat{\mu}_t + \varphi_t) dt + w\left(0, \Omega_{\tau_{i+1}}, \varphi_{\tau_{i+1}}\right) \,\middle|\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}\right] \\
&= \mathbb{E}\left[\int_{0}^{\tau_{i+1}-\tau_i} (\hat{\mu}_t + \varphi_t) dt + w\left(0, \Omega_{\tau_{i+1}-\tau_i}, \varphi_{\tau_{i+1}-\tau_i}\right) \,\middle|\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i}\right] \\
&= \mathbb{E}\left[\int_{0}^{\tau} (\hat{\mu}_t + \varphi_t) dt + w\left(0, \Omega_{\tau}, \varphi_{\tau}\right) \,\middle|(\hat{\mu}_0, \Omega_0, \varphi_0) = (\hat{\mu}_{\tau_i}, \Omega_{\tau_i}, \varphi_{\tau_i})\right]
\end{aligned}$$

where in the second step we split the time between two intervals $[\tau_i, \tau_{i+1}]$ and $[\tau_{i+1}, \infty]$ and use the strong Markov property of our process and the firms policy function, in the third step we substitute the definition of w, in the fourth step we transform the time dimension, and in the fifth step we define $\tau = \tau_{i+1} - \tau_i$, which is equal to $\tau = \inf\{t : |\hat{\mu}_t| \geq \bar{\mu}(\Omega_t)\}$. We arrive to:

$$w(\hat{\mu}_0, \Omega_0, \varphi_0) = \mathbb{E}\left[\int_0^{\tau} (\hat{\mu}_t + \varphi_t) dt + w\left(0, \Omega_{\tau}, \varphi_{\tau}\right)\right]$$

5. **Area under the impulse-response**

Define $F_t(\hat{\mu}, \Omega, \varphi)$ as the cross sectional density over $(\hat{\mu}, \Omega, \varphi)$ in period t after the aggregate shocks and $F_{t0}(\hat{\mu}, \Omega, \varphi | \hat{\mu}_0, \Omega_0)$ the transition probability with initial conditions $(\hat{\mu}_0, \Omega_0, \varphi_0)$. From the definition of $\mathcal{M}(\delta, \alpha, \kappa)$, we have that

$$\begin{aligned}
\mathcal{M}(\delta, \alpha, \kappa) &\equiv -\int_0^{\infty} \hat{Y}_t dt \\
&= -\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega} (\hat{\mu}_t + \varphi_t)\, dF_t(\hat{\mu}_t, \Omega_t, \varphi_t)\right] dt \\
&= -\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega, \varphi} (\hat{\mu}_t + \varphi_t)\, dF_{t0}(\hat{\mu}_t, \Omega_t, \varphi_t | \hat{\mu}_0, \Omega_0, \varphi_0) dG(\hat{\mu}_0, \Omega_0, \varphi_0)\right] dt \\
&= -\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega, \varphi} (\hat{\mu}_t + \varphi_t)\, dF_{t0}(\hat{\mu}_t, \Omega_t, \varphi_t | \hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0) dG(\hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0)\right] dt \\
&= -\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega, \varphi} (\hat{\mu}_t + \varphi_t)\, dF_{t0}(\hat{\mu}_t, \Omega_t, \varphi_t | \hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0) \phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma\Omega_0}\right) d\varphi_0 dF(\hat{\mu}_{-1}, \Omega_{-1})\right] dt \\
&= -\int_{\hat{\mu}, \Omega} \left[\int_0^{\infty} (\hat{\mu}_t + \varphi_t)\, dF_{t0}(\hat{\mu}_t, \Omega_t, \varphi_t | \hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0) dt\right] \phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma(\Omega_{-1} + \kappa\bar{\Omega})}\right) d\varphi_0 dF(\hat{\mu}_{-1}, \Omega_{-1}) \\
&= -\int_{\hat{\mu}, \Omega, \varphi} \left[\int_0^{\infty} \mathbb{E}\left[\hat{\mu}_t + \varphi_t |\, \hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0\right] dt\right] \phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma(\Omega_{-1} + \kappa\bar{\Omega})}\right) d\varphi_0 dF(\hat{\mu}_{-1}, \Omega_{-1}) \\
&= -\int_{\hat{\mu}, \Omega} \int_{\varphi} \mathbb{E}\left[\int_0^{\infty} (\hat{\mu}_t + \varphi_t) dt \middle| \hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0\right] \phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma(\Omega_{-1} + \kappa\bar{\Omega})}\right) d\varphi_0 dF(\hat{\mu}_{-1}, \Omega_{-1}) \\
&= -\int_{\hat{\mu}, \Omega} \int_{\varphi} w(\hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0) \phi\left(\frac{\varphi_0 + (1-\alpha)\delta}{\gamma(\Omega_{-1} + \kappa\bar{\Omega})}\right) d\varphi_0 dF(\hat{\mu}_{-1}, \Omega_{-1})
\end{aligned}$$

where in the second step we use our result that the output deviation at t is equal to the average of markup gap estimates plus forecast errors across firms at each time t, in the third step we factor the distribution at t as the transition probability times the initial distribution after the aggregate shocks, in the fourth step we substitute the initial conditions, in the fifth step we substitute the initial distribution G with the steady state distribution evaluated at the initial conditions times the distribution of forecast errors, in the sixth step we exchange the integrals between time and states, in the seventh step we write in terms of expectations, in the eighth step we exchange the expectation and integral operators, and and in the last step we use the definition of w

$$w(\hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0) = \mathbb{E}\left[\int_0^{\infty} (\hat{\mu}_t + \varphi_t) dt \middle| \hat{\mu}_{-1} - \alpha\delta, \Omega_{-1} + \kappa\bar{\Omega}, \varphi_0\right]$$

□

B Appendix: Model Computation in Discrete Time

The model is solved numerically as a discrete time version of the continuous time model described in the text.

Firm Problem We compute the firms policy function in the steady state solving the firms problem given by

$$
\begin{aligned}
V(\hat{\mu}_-, \Sigma_-) &= \mathbb{E}\left[\max_{c,nc}\{V^c(\Sigma), V^{nc}(\hat{\mu}, \Sigma)\}\right] \qquad \text{(B.46)}\\
V^{nc}(\hat{\mu}, \Sigma) &= -\hat{\mu}^2 + \beta V(\hat{\mu}, \Sigma)\\
V^c(\Sigma) &= \max_x \quad -\theta - x^2 + \beta V(x, \Sigma)\\
\hat{\mu} &= \hat{\mu}_- + \frac{\Sigma}{\sqrt{\Sigma + \gamma^2}}\epsilon\\
\Sigma &= \frac{\gamma^2}{\Sigma_- + \gamma^2}\Sigma_- + \sigma_u^2 J\\
\epsilon &\sim \mathcal{N}(0,1) \qquad J = \begin{cases} 1 & \text{with prob. } p\\ 0 & \text{with prob. } 1-p \end{cases}
\end{aligned}
$$

Value and policy function We approximate a discrete version of the firm value function with 3rd order splines and solve it with iterative and colocation methods.

Steady state To compute the ergodic steady state we apply a histogram approach. Given a discrete grid over $Z = (\hat{\mu}, \Sigma)$, we compute the transition probability $P_{Z'|Z}$ over this grid, and and recover the ergodic distribution as the eigenvector with unit eigenvalue.

Impulse-Reponse To compute the impulse-response to a monetary shock, we use the steady-state policies, since as Álvarez and Lippi (2014) have shown, general equilibrium effects are small. To compute the transition, we compute the transition dynamics over $Z' = (\hat{\mu}, \Sigma, \varphi)$ and compute iteratively the distribution of firms over the grid Z'.

We compute average forecast errors and average markup gap estimates as

$$
ME_t = \sum_{Z'} \hat{\mu}(Z')n_t(Z') \quad ; \quad FE_t = \sum_{Z'} \varphi(\mu)(Z')n_t(Z') \qquad \text{(B.47)}
$$

where $n_t(Z')$ is the distribution at time t. Then the total output effects are given by:

$$
Y_t = -ME_t - FE_t
$$

Uncertainty and Pass-Through We assume the following process for money shocks $\log(M_t) = \log(M_{t-1}) + \sigma_m \epsilon_t$ for σ_m small. Since money shocks are a martingale and σ_m is small, steady state policies are a good approximation of the policies that take into account general equilibrium effects (see Golosov and Lucas (2007) and Álvarez and Lippi (2014)). Thus, we solve steady state policies and simulate a panel of firms. Then we keep the objects needed for the regression in Section 6.