Firm Uncertainty Cycles and the Propagation of Nominal Shocks*

Isaac Baley[†] **Julio A. Blanco** [‡]

February 21, 2018

Abstract

We develop a framework to study the impact of idiosyncratic uncertainty on aggregate economic outcomes. Agents learn about idiosyncratic characteristics, which receive infrequent, large, and persistent shocks. In this environment, *idiosyncratic uncertainty* moves in cycles, fluctuating between periods of high and low uncertainty; with additional fixed adjustment costs, the frequency and size of agents' actions also fluctuates in cycles. We apply our framework to study pricing behavior and the propagation of nominal shocks. We show, analytically and quantitatively, that idiosyncratic uncertainty cycles amplify the real effects of nominal shocks, by generating cross-sectional dispersion in firms' adjustment frequencies and in learning speeds.

JEL: D80, E30, E50

Keywords: Adjustment costs, menu costs, fat-tailed shocks, firm-level uncertainty, price-setting, information frictions, monetary policy, forecast errors.

*We are especially thankful to Virgiliu Midrigan and Laura Veldkamp for their advice and to the referees and the Editor for their constructive comments. We also thank Fernando Álvarez, Rudi Bachmann, Anmol Bhandari, Jarda Borovička, Katka Borovičková, Olivier Coibion, Mark Gertler, Julian di Giovanni, Christian Hellwig, Ricardo Lagos, John Leahy, Francesco Lippi, Robert E. Lucas, Rody Manuelli, 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 NYU, NYU Stern, Princeton, Washington University St. Louis, U Toronto, EIEF, CREI, Pompeu Fabra, BIS, Singapore Management U, Carnegie Mellon, UC Davis, U Melbourne, U Sydney, ITAM, Oxford, UAB, U Mannheim, U Michigan, U Minnesota, 4th Ifo Conference 2013, Midwest Economics Association 2013, SED 2013, ASSA 2015, SITE 2015, Econometric Society 2015, 40 Simposio de la AEE, Barcelona GSE Summer Forum 2016, Federal Reserve Board, St. Louis Fed, Banco de México, Banque de France, and Banco do Brasil for very useful comments and suggestions. Isaac Baley gratefully acknowledges the support of the Spanish Ministry of the Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563) and Project ECO2015-67655-P. 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 and marketing campaigns get developed, unfamiliar markets are targeted, workers are replaced, and supply chains get disrupted. These idiosyncratic changes are recurrent, large, and persistent, and under many scenarios, firms do not have all the information needed to assess their impact. This lack of perfect knowledge generates uncertainty that affects firms' actions. Moreover, responding to the new conditions may entail large adjustment costs for firms. In this context, many questions arise naturally: How do firms process and respond to uncertain changes in their environment? Is it possible to identify the degree of frictions from the data? Does firm-level uncertainty matter for aggregate economic outcomes?

To address these questions, we develop a general framework where agents have imperfect information about persistent idiosyncratic conditions that change infrequently and by large amounts—fat-tailed risk—and must pay a fixed cost to make choices. The model is very general and can be applied to a wide array of settings. For example, it can be used to analyze portfolio choices subject to adjustment fees and uncertain trends in asset returns, or workers' occupational choices subject to mobility costs and uncertain productivity growth. We apply the framework to study firms' pricing decisions and the impact of firm-level uncertainty on the propagation of aggregate nominal shocks.

We model price-setting firms that face menu costs to adjust their prices and imperfect information about their idiosyncratic productivity. We assume that firms receive persistent shocks to their productivity that are fat-tailed, where only the timing of the shocks is known, but not the exact realization. For example, a firm knows a worker or its management has been replaced, but it does not know how productive it will be under the new regime. Firms use Bayes' law to estimate their persistent productivity as in Jovanovic (1979), and we call the conditional variance of the estimates firm uncertainty. In this setup, firm-level uncertainty cycles arise. Uncertainty cycles are defined as recurrent patterns in which firm uncertainty spikes up when a fat-tailed productivity shock hits, and then fades with learning until the arrival of the next shock, when uncertainty jumps again. Due to fixed adjustment costs, the pricing policy takes the form of an inaction region that moves with uncertainty; as a consequence, the frequency and the size of price changes also fluctuate in cycles. With a continuum of ex-ante identical firms, the uncertainty cycles endogenously generate dispersion in adjustment frequency as firms churn between high and low levels of uncertainty: high uncertainty firms change their prices more often than low uncertainty firms. Additionally, the uncertainty cycles generate dispersion in firms' learning dynamics: high uncertainty firms incorporate new information more quickly into their forecasts than low uncertainty firms.

Our key prediction is that idiosyncratic uncertainty cycles amplify the real effects of nominal shocks. This result is robust to a variety of specifications, including scenarios with partial knowledge about the size of the nominal shock and the introduction of aggregate uncertainty shocks. Amplification arises from various forces generated by the uncertainty cycles, two of which are new in the literature: (i) dispersion in times until the first price adjustment, (ii) dispersion in learning dynamics, and (iii) a positive correlation between adjustment frequency and the degree of selection effect. When we calibrate the model to match a variety of micro-price statistics, we obtain an output response that is twice as large as the benchmark in Golosov and Lucas (2007), where pricing decisions are identical across firms. Moreover, the output effects can be up to seven times larger if the monetary shock is only partially observed, in which case firms make forecast errors about the monetary shock. All these results highlight the role of firm-level uncertainty, especially its cross-sectional distribution, in assessing the effectiveness of monetary policy.

Our analysis proceeds in three steps. First, we show how uncertainty cycles arise endogenously from firms' learning processes in an environment with fat-tailed risk. The departure from the Gaussian world entails important computational challenges, but we keep the Bayesian learning tractable by assuming that the timing of shocks is known. Importantly, both the information friction and the fat-tailed shocks are needed to generate heterogeneity in uncertainty in the steady state. Each ingredient on its own is not enough, however. With the information friction alone, all firms eventually track their productivity with equal and constant precision. Álvarez, Lippi and Paciello (2011) study that case and conclude that the economy collapses to the standard menu cost model of Golosov and Lucas (2007) with homogeneous firms and large monetary neutrality. The novelty in our setup is that such stabilization is prevented by the fat-tailed shocks, and heterogeneity persists in steady state. With fat-tailed shocks alone, firms face an almost constant probability of adjustment as in Gertler and Leahy (2008) or Midrigan (2011). By incorporating the information friction, our model captures several features from the micro-pricing data that cannot be explained by those benchmarks, and reveals new mechanisms that result in larger monetary non-neutrality.

Uncertainty cycles are empirically relevant. Bachmann, Elstner and Sims (2013) and Bachmann, Elstner and Hristov (2017b) find substantial cross-sectional heterogeneity and time-variation in measures of firm-idiosyncratic uncertainty using survey data for German firms. Similar evidence is documented for the US in Senga (2016) by merging survey data from the I/B/E/S and Compustat. The evidence reveals persistent differences in the degree of firm-level uncertainty, not only in the cross-section, but also across time within the same firm. Regarding fat-tailed risk affecting firms, Klenow and Malin (2011) and Alvarez, Le Bihan and Lippi (2016) document positive excess kurtosis in the price change distribution for US and French CPI data, respectively. Further evidence for underlying fat-tailed risks is found in many firm variables beyond the price change distribution. Davis and Haltiwanger (1992) provide clear evidence of leptokurtic changes in employment in US Census data. As further suggestive evidence, we document fat-tailed distributions for the growth rates of profits, employment, sales, and capital in Compustat firms between 1980 and 2015.[1]

The second step in the analysis focuses on the effects that uncertainty cycles have on firm pricing decisions. The answer is not obvious, as there are two opposing effects of uncertainty. More uncertain firms have wider inaction regions which makes prices less flexible—"option value effect" in Dixit (1991)—but they also change their beliefs in a more volatile way, which makes prices more flexible—"volatility effect". We show analytically that the latter effect dominates; thus higher uncertainty yields higher adjustment frequency and bigger price changes. As a result of the positive relationship

[1]These variables are clearly endogenous objects and not structural shocks. Nevertheless, in a large set of models, these variables are related to firms' productivity. See Online Appendix Section A.1.

between uncertainty, frequency, and size, the firm uncertainty cycles imply adjustment cycles: periods of frequent adjustment and large price changes alternate with periods of infrequent adjustment and small price changes, i.e. price changes get clustered instead of evenly spread across time. This implies that the probability to adjust conditional on no past adjustment—the hazard rate—falls with time. Moreover, the magnitude of the information frictions pins down the hazard rate's slope. To understand these results, note that the inaction regions refer to productivity estimates and not the true realizations. This makes a difference because after a firm takes action, its judgment might turn out to be wrong, leading it to take further action again very soon. This contrasts sharply with standard menu cost models, where the probability of adjustment right after a price change is tiny and increases over time.

There is empirical evidence that supports the predictions of our price-setting model. Bachmann, Born, Elstner and Grimme (2017a) document a positive correlation between firm-level uncertainty and price adjustment for German firms. Campbell and Eden (2014) document that price changes in the retail sector are more extreme and dispersed for recently changed prices compared to older prices. Finally, decreasing hazard rates are documented for various datasets, countries and time periods.[2]

In the third and last part of the analysis, we study the effects of uncertainty cycles in the propagation of monetary shocks. For this purpose, we consider a Bewley–type economy with a continuum of ex-ante identical firms that face menu costs and idiosyncratic uncertainty cycles. This economy features a non-degenerate steady state distribution where a Pareto principle operates: a small proportion of firms has high uncertainty but is responsible for the majority of belief updates and price changes, while the majority of firms has low uncertainty and hardly contributes to belief updates and price changes. In this economy we study the effect of a small, unanticipated, and permanent increase in the money supply. We demonstrate analytically that monetary shocks have larger effects on output than in an alternative economy with a representative firm, conditional on the same average adjustment frequency. We also show that these effects are quantitatively important in a calibrated version of the economy that matches micro-price statistics.[3]

Let us now explain the sources of amplification as a result of uncertainty cycles. The first amplification mechanism arises due to dispersion of times until the first price adjustment following a monetary shock. To understand its logic, it is key to recognize that a firm's *first* price change after a monetary shock incorporates that shock into its 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. High uncertainty firms adjust quickly to incorporate the monetary shock (they are already adjusting for idiosyncratic reasons), whereas low uncertainty firms take a long time to adjust. Through a Jensen's inequality, this dispersion allows to match average price statistics while increasing aggregate price rigidity. This mechanism is explored in Carvalho (2006), Nakamura and Steinsson (2010), Bouakez, Cardia and Ruge-Murcia (2014), Carvalho and Schwartzman (2015) and Álvarez, Lippi and Paciello (2016). While in those pa-

[2]See Dhyne *et al.* (2006), Nakamura and Steinsson (2008), Eden and Jaremski (2009), Vavra (2010), Cortés, Murillo and Ramos-Francia (2012), Campbell and Eden (2014), and Argente and Yeh (2015).

[3]Uncertainty cycles have consequences for the interpretation of price-change statistics. Average price-change statistics reflect more intensively the pricing behavior of highly uncertain firms because they are more prone to adjust than the average. We explain how to correct for this bias and show that, once corrected, the aggregate price level is stickier than suggested by average frequency.

pers heterogeneity is exogenous and fixed, in our model heterogeneity arises endogenously in ex-ante identical firms that churn between high and low levels of uncertainty.

This brings us to the second and more subtle amplification mechanism. The uncertainty cycles basically generate two sets of firms. The first group consists of low frequency adjusters that change their prices *primarily* due to the arrival of infrequent shocks as in Gertler and Leahy (2008); their pricing behavior features small selection effects and are responsible for most of the real effects of nominal shocks. The second group consists of high frequency adjusters that *primarily* change their prices due to the diffusion process as in Golosov and Lucas (2007); their pricing behavior features strong selection effects that dampen the real effects of nominal shocks. This means that uncertainty cycles generate a positive correlation between the strength of the selection effects and the frequency of adjustment: the low frequency adjusters are also those that does not respond much to monetary shocks. We show that this correlation, which is novel force in these models, further amplifies the real effects of monetary shocks.

The third amplification mechanism, exclusive to our learning model with uncertainty cycles, refers to the dispersion in the learning speed across firms: high uncertainty firms optimally put a larger Bayesian weight on new information compared to low uncertainty firms, and consequently, they learn faster than low uncertainty firms. To highlight the workings of this mechanism, we assume that the monetary shock is only partially observed by firms and that they filter it out using the same learning technology they use to estimate their idiosyncratic productivity. To bring discipline to the observability of the monetary shock, we use evidence from survey forecast data in Coibion and Gorodnichenko (2012). With partial observability of the monetary shock, the first adjustment does not fully take care of incorporating it into prices—passthrough is incomplete—and there will be forecast errors that do not cancel out in the aggregate. We demonstrate that the existence of average forecast errors amplifies real output effects, and by Jensen's inequality once again, that forecast errors' persistence is increasing in the dispersion of Bayesian updating weights. Therefore, uncertainty cycles generate highly persistent average forecast errors that extend the effects of monetary shocks.

Finally, as a way to further understand the learning mechanism, we consider a situation where all firms experience a synchronized uncertainty shock at the same time of the monetary shock. We discover that this *aggregate* uncertainty shock increases all firms' learning speed and the persistence of average forecast errors falls. This implies that monetary shocks have smaller real effects when *aggregate* uncertainty is large. Empirical support for this channel can be found in Aastveit, Natvik and Sola (2013), Vavra (2014), Caggiano, Castelnuovo and Nodari (2014), and Castelnuovo and Pellegrino 2018, who find weaker effects of monetary policy when economic uncertainty is high, and in Coibion and Gorodnichenko (2015), who provide evidence that the degree of information rigidity responds to changing economic conditions. The joint dynamics of uncertainty, price-setting, and forecast errors implied by our model, and disciplined with micro data, provide a framework to interpret this evidence.

2 Pricing with uncertainty cycles

We combine an inaction problem with a signal extraction problem. Its structure is very general and it is easily extendable to a variety of environments that involve non-convex adjustment costs and imperfect information about fat-tailed idiosyncratic shocks.

2.1 Environment

Consider a profit maximizing firm that chooses the price at which to sell its product, subject to idiosyncratic productivity shocks. It must pay a menu cost θ in units of product every time it changes the price. We assume that in the absence of the menu cost, the firm would like to set a price that makes its markup—price over marginal cost—constant. The productivity shocks are not perfectly observed; only noisy signals are available to the firm.[4] It 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 losses. Let μ_t be the markup gap, defined as the log difference between the current markup and the frictionless markup. Firms incur an instantaneous quadratic loss as the markup gap moves away from zero:

$$\Pi(\mu_t) = -B\mu_t^2, \qquad B > 0. \tag{1}$$

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

Markup gap process. Firms make inferences about their productivity, but when the price is fixed, the markup gap follows productivity. Therefore, it is equivalent and convenient to work directly with markup gaps. We assume μ_t follows a jump-diffusion process

$$d\mu_t = \sigma_f dW_t + \sigma_u u_t dQ_t, \tag{2}$$

where W_t is a Wiener process, $u_t Q_t$ is a compound Poisson process with 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 specification nests two benchmarks in the literature: small *frequent* shocks modeled as the Wiener process W_t with small volatility σ_f, as in Golosov and Lucas (2007), and large *infrequent* shocks modeled through the Poisson process Q_t with large volatility σ_u, as in Gertler and Leahy (2008) and Midrigan (2011).[5]

[4] 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.

[5] Alternatively, markup gap shocks can be interpreted as shocks to the frictionless markup, e.g. shocks to the demand elasticity in a CES framework. See Online Appendix Section A.2 for details.

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

$$ds_t = \mu_t dt + \gamma dZ_t, \tag{3}$$

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 μ_t enters as the drift of the signal. This representation makes the filtering problem tractable as the signal is continuous.[6]

Information set. We assume that a firm knows if they receive an infrequent shock but not the size of the innovation u_t. Thus the information set at time t is given by the σ-algebra generated by the history of signals s and realizations of Q:

$$I_t = \sigma\{s_r, Q_r;\ r \leq t\}. \tag{4}$$

Since Poisson innovations u_t are not observed but have conditional mean of zero, firms know that infrequent shocks could push their markups either upwards or downwards, but in expectation they have no effect. This assumption allows us to match the price change distribution's symmetry around zero, and largely simplifies the exposition. But it is not crucial.[7]

2.2 Filtering problem

Firms' learning technology. Firms make estimates in a Bayesian way by optimally weighing new information from signals against old information from prior estimates. This is a passive learning technology in the sense that firms process the information that is available to them, but they cannot take 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 by experimenting with price changes.

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 establishes the laws of motion for markup gap estimates and uncertainty keeping the finite-state properties of the Gaussian model by representing the posterior distribution $\mu_t | \mathcal{I}_t$ as a function of mean and variance. Our contribution extends the Kalman–Bucy filter beyond the standard assumption of Brownian motion innovations. To our knowledge, this is a novel result in the filtering literature. All proofs are provided in the Appendix.

[6] See Chapter 6 in Øksendal (2007) and the Online Appendix Sections B.1 and B.2 for details.

[7] In the Online Appendix Section B.3 we extend the results to consider a positive or negative mean and provide further technical discussion for interested readers.

Proposition 1 (**Filtering equations**). *Let the markup gap and the signal evolve according to (2) and (3), and consider the information set in (4). Then the posterior distribution of markup gaps is Gaussian $\mu_t|\mathcal{I}_t \sim \mathcal{N}(\hat{\mu}_t, \gamma\Omega_t)$, where $(\hat{\mu}_t, \Omega_t)$ evolve as follows:*

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

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

$\hat{Z}_t$ 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 firm's information set, and it is independent of dQ_t.

Uncertainty increases the volatility of estimates. 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 (5). Consider a small period of time Δ. The markup gap estimate at time $t+\Delta$ is given by the convex combination of the prior estimate $\hat{\mu}_t$ and the signal change $s_t - s_{t-\Delta}$:

$$\hat{\mu}_{t+\Delta}\Delta = \underbrace{\frac{\gamma}{\Omega_t\Delta + \gamma}}_{\text{weight on prior estimate}} \hat{\mu}_t\Delta + \underbrace{\left(1 - \frac{\gamma}{\Omega_t\Delta + \gamma}\right)}_{\text{weight on signal}} (s_t - s_{t-\Delta}). \tag{7}$$

Due to Bayesian updating, when uncertainty is high, estimates optimally put more weight on signals instead of the prior. Learning is faster, but it also brings more white noise into the estimation. Estimates become more volatile with high uncertainty. This effect is key in our discussion of price responsiveness to monetary shocks.

Idiosyncratic uncertainty cycles. Equation (6) shows that uncertainty has a deterministic and a stochastic component. In the absence of fat-tailed shocks ($\lambda = 0$), uncertainty follows a deterministic path which converges to the constant volatility of the continuous shocks σ_f. With fat-tailed shocks ($\lambda > 0$), the time series profile of uncertainty features a saw-toothed profile that never stabilizes: uncertainty jumps up after the arrival of an infrequent shock and then decreases deterministically until the arrival of the following one. Although uncertainty never settles down, there is a "long-run" level Ω^* such that its expected change is zero, $\mathbb{E}\left[d\Omega_t \middle| \mathcal{I}_t\right] = 0$. It is equal to $\Omega^* \equiv (\sigma_f^2 + \lambda\sigma_u^2)^{\frac{1}{2}}$. The ratio of current to long-run uncertainty Ω_t/Ω^* appears in decision rules and price statistics and it is characterized in Proposition 8 as a function of price statistics.

With the filtering problem at hand, the following section derives the price adjustment decision.

2.3 Decision rules

Sequential problem. Let $\{\tau_i\}_{i=1}^{\infty}$ be the series of dates where the firm adjusts its 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 state's law of motion, and the filtration $\{\mathcal{I}_t\}_{t=0}^{\infty}$, the sequential problem 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 in (8) is solved recursively as a stopping time problem using the Principle of Optimality, see Øksendal (2007) and Stokey (2009) for details. The following points are formalized in Proposition 2. First, the markup gap μ_t can be substituted for its estimate $\hat{\mu}_t$ without altering the optimal policy, as the difference in payoffs is a constant that arises from firms' inability to perfectly learn the true realizations. Second, since we are working in a Gaussian/quadratic framework, the firm's state is fully characterized by its markup gap estimate $\hat{\mu}$ and the uncertainty attached to that estimate Ω, i.e. higher moments are not needed. Third, given its state $(\hat{\mu}_t, \Omega_t)$, the policy consists of (i) a stopping time τ, which is a measurable function with respect to the filtration, and (ii) the new markup gap μ'.

Proposition 2 (**Stopping time problem**). *Let $(\hat{\mu}_0, \Omega_0)$ be the firm's 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 (5) and (6) .

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. The inaction region is two-dimensional because the firm has two states. Let $\bar{\mu}(\Omega)$ denote the inaction region's border as a function of uncertainty. The inaction region is described by the set $\mathcal{R} = \{(\hat{\mu}, \Omega) : |\hat{\mu}| \leq \bar{\mu}(\Omega)\}$. Its symmetry around zero is inherited from the specification of the stochastic process, the quadratic profits, and zero inflation. For the same reasons, the reset markup gap is equal to zero, i.e. $\hat{\mu}' = 0$. This means that, upon adjustment, the firm chooses a price that it thinks will bring its markup to the frictionless level, but its judgement might be wrong.

This inaction problem is non-standard because 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), the standard smooth pasting condition for both states.[8] Proposition 3 formalizes these points.

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

[8]In the Online Appendix, Section C derives the infinitesimal generator and its adjoint operator; Section D verifies that the conditions in that Theorem hold in our problem; and Section E.3 verifies numerically that the smooth pasting conditions for $\hat{\mu}$ and Ω are valid.

conditions, 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:

1. *In the interior of the inaction region, 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, ϕ satisfies the value matching condition, which sets the value of adjusting equal to the value of not adjusting:*

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

3. *At the border of the inaction region, ϕ satisfies two smooth pasting conditions, one per state:*

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

The passive learning process of our model implies the lack of interaction terms between uncertainty and markup estimates, since firms cannot change the information flow. Proposition 4 analytically characterizes the inaction region's border $\bar{\mu}(\Omega)$ using the three conditions above. The proof uses a Taylor expansion of the value function.[9]

Proposition 4 (Inaction region). *For small r and $\bar{\theta}$, 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 the inaction region with respect to uncertainty is equal to

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

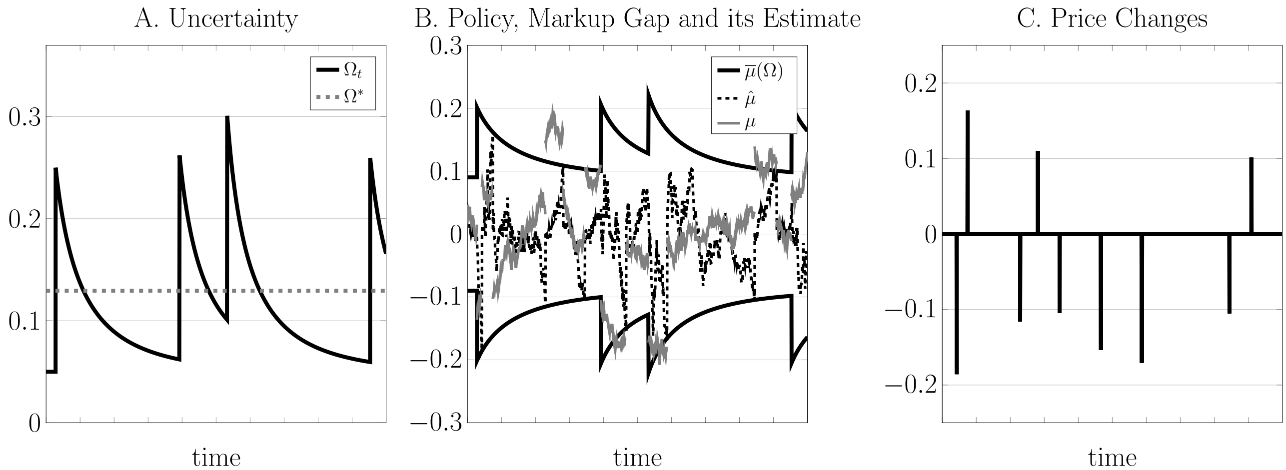
Uncertainty widens inaction region. The numerator of the inaction region in equation (13) is increasing in uncertainty and captures the well known *option value effect.* As a result of uncertainty dynamics, the option value is time varying. This mechanism is reflected in the factor $\mathcal{L}^{\bar{\mu}}(\Omega)$ that amplifies or dampens the option value effect depending on the ratio of current uncertainty to long-run uncertainty Ω/Ω^*. When current uncertainty is high $\left(\frac{\Omega}{\Omega^*} > 1\right)$, it is expected to decrease ($\mathbb{E}[d\Omega] < 0$) and future option values also decrease. This feeds back into the current inaction region, shrinking it as $\mathcal{L}^{\bar{\mu}}(\Omega) > 0$. The total effect of uncertainty on the inaction region depends on the ratio of the menu cost and the signal noise. With small menu costs θ and large signal noise γ, the inaction region is

[9] Section E 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 section.

increasing in uncertainty, with an elasticity of the inaction region with respect to uncertainty $\mathcal{E}(\Omega)$ close to 1/2. The critical result is that the elasticity is less than unity for all possible parametrizations.

Figure I shows one firm's realization. Panel A shows the evolution of uncertainty, which follows a saw-toothed profile. Panel B plots the true markup gap, the markup gap estimate and the inaction region. As we can see in the figure, the markup gap estimate is always within the Ss bands, but this is not the case for the true markup gap. Additionally, the inaction region follows uncertainty's profile as it is increasing in uncertainty given the calibration. Finally, Panel C shows the magnitude of price changes, which are triggered when the markup gap estimate touches the border of the inaction region. Notice that price changes are clustered over time, i.e. there are recurrent periods with high adjustment frequency followed by periods of low adjustment frequency. Moreover, as the width of inaction regions decreases over time, the size of price changes also falls.

Figure I – Sample Paths For One Firm



Panel A: Uncertainty (solid line) and long-run uncertainty (dotted line). Panel B: True markup gap (gray solid line), Markup gap estimate (black dotted line) and inaction region (black soli line). Panel C: Magnitude of price changes. This figure simulates one realization of the stochastic processes using the finite difference method and uses the analytical approximation of the inaction region.

Without fat-tailed shocks, uncertainty converges to a constant, i.e., $\Omega \to \sigma_f$, the inaction region is constant, and there is no dispersion in the size of price changes. This makes evident that both the fat-tailed shocks and the information friction are key to generate the cross-sectional variation in price setting that arises from uncertainty cycles.[10] The next section derives expressions for price statistics that make evident the impact of uncertainty cycles on pricing decisions.

2.4 Uncertainty and conditional price statistics

How does uncertainty affect the adjustment frequency? There are two opposing forces. Uncertainty increases estimate volatility, raising the probability of hitting the bands and adjusting the price. To save on menu costs, the inaction region widens, reducing the adjustment probability.

[10]The case without fat-tailed shocks is analyzed in Álvarez, Lippi and Paciello (2011). That paper shows that the model collapses to that of Golosov and Lucas (2007) where all firms have the same inaction region.

The first effect dominates because the elasticity of the inaction region with respect to uncertainty is lower than one, and higher uncertainty increases adjustment frequency. This result holds under our assumptions of a quadratic payoff function and martingale processes. Proposition 5 formalizes this result.

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:*

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

If the elasticity of the inaction region with respect to uncertainty is less than one, then the expected time between price changes $\mathbb{E}[\tau \big| 0, \Omega]$ is a decreasing and convex function of uncertainty.

The conditional expected time between price changes has two terms. The first term $\frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}$ states that expected time for adjustment is smaller when the markup estimate is closer to the inaction border. This term is decreasing in uncertainty with an elasticity larger than one in absolute value. The second term $\mathcal{L}^\tau(\Omega)$ amplifies or dampens the first effect depending on the level of uncertainty, and it has an elasticity of one. Uncertainty's overall effect on expected adjustment time is negative: more uncertain firms change their prices more frequently than less uncertain firms.

A new insight from our model is that inaction regions refer to markup estimates and not the true realizations. Firms adjusts their prices so that their expected markup gap becomes zero, but this expectation is surrounded by uncertainty. After a price change, a high uncertainty firm is very likely to have made a wrong adjustment which leads to a new price change. As it learns its new productivity level, the likelihood of further price changes falls. This contrasts sharply with a standard menu cost model where the probability of adjustment right after a price change is tiny and increases over time, or with a Calvo model where the probability of adjustment is always the same. The hazard rate is the adequate statistic to measure these effects and also distinguish across models, as it is a dynamic measure of adjustment frequency.

The conditional hazard rate, denoted by $h_\tau(\Omega)$, is the probability of adjusting τ periods after the last price change, and it is conditional on uncertainty Ω. It depends on the expected width of the inaction region relative to the expected level of uncertainty τ periods ahead. Proposition 6 characterizes the conditional hazard rate by making two simplifications: it assumes a constant inaction region and shuts down future infrequent shocks.[11]

Proposition 6 (**Conditional Hazard Rate**). *Without loss of generality, assume a firm's last price change occurred at $t = 0$ and it has an initial level of uncertainty $\Omega_0 > \sigma_f$. Assume the inaction region is constant $\bar{\mu}_0$ and no further infrequent shocks are expected ($\lambda = 0$). Denote derivatives with respect to τ with a prime ($h'_\tau \equiv \partial h / \partial \tau$).*

[11] Online Appendix E.7 computes the exact numerical hazard rate and checks the validity of these assumptions.

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

$$\mathcal{V}_\tau(\Omega_0) = \sigma_f^2\tau + \gamma(\Omega_0 - \Omega_\tau), \quad \textit{where} \quad \Omega_\tau = \sigma_f \left(\frac{\frac{\Omega_0}{\sigma_f} + \coth\left(\frac{\sigma_f}{\gamma}\tau\right)}{1 + \frac{\Omega_0}{\sigma_f}\coth\left(\frac{\sigma_f}{\gamma}\tau\right)} \right) \tag{16}$$

and it is an increasing and concave function of duration τ and initial uncertainty Ω_0.

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

$$h_\tau(\Omega_0) = \frac{\pi^2}{8} \underbrace{\Psi\left(\frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)}_{\textit{increasing in } \tau} \underbrace{\frac{\mathcal{V}'_\tau(\Omega_0)}{\bar{\mu}_0^2}}_{\textit{decreasing in } \tau}, \tag{17}$$

where $\Psi(\cdot)$ is the increasing hazard rate for a case with fixed uncertainty $\Omega_0 = \Omega_\tau = \sigma_f$ derived in Kolkiewicz (2002), with $\Psi(0) = 0$, $\Psi'(x) > 0$, $\lim_{x\to\infty}\Psi(x) = 1$, first convex then concave.

3. *There exists a date $\tau^*(\Omega_0)$ such that $h'_\tau(\Omega_0) < 0$ for $\tau > \tau^*(\Omega_0)$, and $\tau^*(\Omega_0)$ is decreasing in Ω_0.*

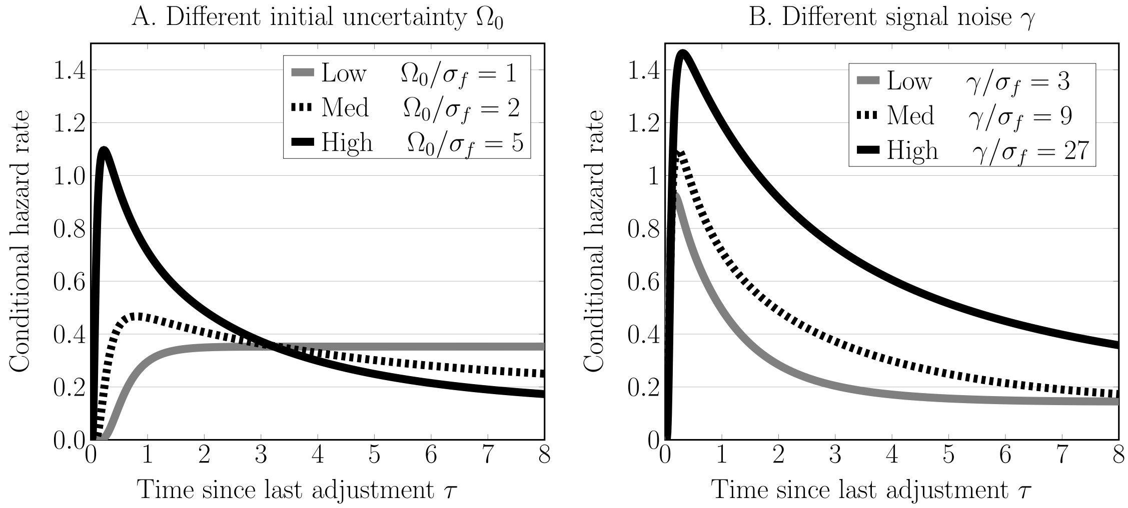
Equation (17) expresses the hazard rate as a product of two factors. The first factor $\Psi(\cdot)$ is responsible for the increasing part close to $\tau = 0$ and resembles the hazard rate of standard menu cost models, as the probability of an additional adjustment right after a price change is very low. The second factor is the change in unconditional variance $\mathcal{V}'_\tau(\Omega_0)$, which is decreasing in τ because of learning. These two opposing forces act upon the slope of the hazard rate and the hazard rate is non-monotonic. The last point in the proposition establishes that there exists a date τ^* after which the hazard is downward sloping, and that this date is shorter the higher the initial uncertainty.

Figure II illustrates the effect of initial uncertainty Ω_0 (Panel A) and signal noise γ (Panel B) on the hazard. Initial uncertainty determines which of the two factors dominates. If initial uncertainty is small, there is not much to learn and the hazard rate is increasing, as in the standard menu cost model. In contrast, if initial uncertainty Ω_0 is large, the learning force dominates and the hazard is decreasing for a larger range of price durations. Signal noise γ determines how fast uncertainty converges to its limit value—independent of γ—and the rate of change in the slope. The larger the information friction (the larger γ), the longer it takes to discover productivity, and the slope decays more slowly. This relationship between γ and the hazard' slope will be used for calibration.

2.5 Aggregation of heterogeneous firms

In this section we explain how to aggregate the conditional price statistics. The goal is to assess whether heterogeneity in pricing behavior has effects in the aggregate and to match the price statistics generated by the model with those in the data. For this purpose, we consider a continuum of ex-ante identical firms that face the pricing problem from the previous sections. Markup shocks and signals are assumed to be *iid* across firms. Independence and stationarity of the controlled process ensure the existence of an ergodic distribution $F(\hat{\mu}, \Omega)$.

Figure II – Conditional Hazard Rate



Panel A. Conditional hazard for three levels of initial uncertainty Ω_0, expressed as multiples of σ_f. Signal noise is fixed at $\gamma = 9\sigma_f$. Panel B. Conditional hazard for three levels of signal noise γ, expressed as multiples of σ_f. Initial uncertainty is fixed at $\Omega_0 = 5\sigma_f$. These are approximated hazard rates with constant inaction regions and without further Poisson shocks. We use a larger σ_f than in the final calibration for illustration purposes.

The aggregate statistics are equal to the weighted average of the conditional statistics, where the weights are given by the distribution of uncertainty *conditional* on price adjustment, denoted with $r(\Omega)$, also known as renewal distribution. This distribution is different from the *unconditional* steady state distribution of uncertainty $h(\Omega)$, which is the uncertainty in the entire cross-section. Proposition 7 shows that the ratio between the renewal and marginal distributions is increasing in uncertainty, i.e. adjusting firms are on average more uncertain than the rest of the population.

Proposition 7 (Uncertainty Distributions). *Assume $\bar{\mu}'(\Omega) > 0$. Then, in a neighborhood around long-run uncertainty Ω^*, the ratio between the renewal and marginal densities of uncertainty are proportional to the inverse of the expected time between adjustments*

$$\frac{r(\Omega)}{h(\Omega)} \propto \frac{1}{\mathbb{E}[\tau|(0,\Omega)]}. \tag{18}$$

As expected time decreases with uncertainty, the ratio $r(\Omega)/h(\Omega)$ increases with uncertainty.

The previous result implies that average price change statistics (computed with the renewal distribution $r(\Omega)$) reflect more intensively the pricing behavior of highly uncertain firms because they are more prone to adjust than the average firm. For instance, if the unconditional hazard rate is decreasing, it is because the renewal distribution puts a large weight on the decreasing hazard rate of high uncertainty firms compared to the increasing hazard rate of low uncertainty firms. The same logic applies to the expected time to adjustment and other price statistics. However, as we show in the following section, the aggregate response to nominal shocks depends on the behavior of the full cross-section (computed with the unconditional distribution of uncertainty $h(\Omega)$). Therefore, uncertainty cycles alter the interpretation of price statistics as a high frequency of adjustment is perfectly compatible with a low aggregate price flexibility. Equation (18) expresses the relationship between these two densities as a function of the expected price duration.

Our second aggregation result, derived in Proposition 8, establishes a link between the long-run uncertainty Ω^*, cross-sectional heterogeneity in uncertainty $\mathbb{E}[\Omega^2]$, adjustment frequency $1/\mathbb{E}[\tau]$, and price change dispersion $\mathbb{V}[\Delta p]$.[12] Its key point is that observable price statistics provide a direct way to recover the heterogeneity in firm level uncertainty.

Proposition 8 (**Uncertainty Heterogeneity and Price Statistics**). *The following relationships between long-run uncertainty, uncertainty dispersion, average price duration, and price change dispersion hold:*

$$\Omega^{*2} = \mathbb{E}[\Omega^2] = \frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}. \tag{19}$$

Holding fixed uncertainty's second moment in the left-hand side, expression (19) establishes a positive link between average price duration and price change dispersion: prices either change often for small amounts or rarely for large amounts. Analogously, 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 duration.

Remarks. We finish this section with a few remarks that emphasize that uncertainty cycles go beyond being a micro-foundation for heterogeneity in price-setting. Moreover, our particular notion of uncertainty shocks (and the learning that follows) plays a specific and important role in generating the type of pricing statistics that we use to discipline model parameters, and this matters in the aggregate. A first alternative model considers an autoregressive process for uncertainty/volatility, as in the menu cost models of Vavra (2014) and Karadi and Reiff (2014). A second alternative model is one in which firms perfectly learn their idiosyncratic state at random times, and at other times they observe noisy signals. This model generates "downward jumps" in uncertainty which contrast with our "upward jumps". We discover that these alternative models produce very different predictions for micro-level price statistics, e.g. increasing re-pricing hazards. Consequently, the micro-price data allow us to distinguish between the type of uncertainty processes faced by firms, as well as the different forms of uncertainty cycles. And this distinction matters for the aggregate results.[13]

3 Aggregate effects of uncertainty cycles

What are the macroeconomic consequences of firm uncertainty cycles and the heterogeneity they generate? We develop a standard general equilibrium framework with monopolistic firms that face the pricing problem with menu costs and uncertainty cycles studied in the previous sections. We use the

[12] This generalizes Proposition 1 in Álvarez, Le Bihan and Lippi (2014) for the case of heterogeneous uncertainty. Blanco (2016) shows that a similar result holds in the case of positive inflation.

[13] In the Online Appendix, Section F.1 compares autoregressive and jump uncertainty processes, while Section F.2 compares upward and downward uncertainty cycles.

model to study the role of firm uncertainty in the propagation of monetary shocks. The model builds on Golosov and Lucas (2007), with the addition of the information friction and fat-tailed shocks.

3.1 General equilibrium model

Time is continuous. There is a representative consumer, a continuum of monopolistic firms, and a monetary authority. We focus on a steady state in which money supply is constant at a level M.

Representative Household The household has the following 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 and the future is discounted at rate $r > 0$:

$$\mathbb{E}_0 \left[\int_0^\infty e^{-rt} \Big(\log C_t - \alpha N_t + \log \frac{M_t}{P_t} \Big) dt \right]. \tag{20}$$

Consumption consists of a CES aggregator as in Woodford (2009), Midrigan (2011), and Álvarez and Lippi (2014). The household has access to complete financial markets. The budget includes labor earnings $E_t N_t$, 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 D_t be the stochastic discount factor; by complete markets, the time-0 Arrow-Debreu budget constraint reads:

$$\mathbb{E}_0 \left[\int_0^\infty D_t \left(P_t C_t + R_t M_t - E_t N_t - \Pi_t \right) dt \right] \leq M_0. \tag{21}$$

The household chooses consumption, labor supply and money holdings to maximize (20) subject to (21). The household's first order conditions establish nominal wages as a proportion of the (constant) money stock $E = rM$.

Monopolistic Firms On the production side, there is a continuum of firms indexed by $z \in [0, 1]$ who operate in a monopolistically competitive market. Each firm maximizes its expected stream of profits, discounted at D_t. It chooses a price and then satisfies all its demand. For every price change, it must pay a menu cost θ. Production uses a linear technology with labor as its only input: producing $y_t(z)$ units requires $l_t(z) = y_t(z) A_t(z)$ units of labor, so that the marginal nominal cost is $A_t(z)E$. We define markups as $\mu_t(z) \equiv \frac{p_t(z)}{A_t(z)E}$. Given the consumer's strategy, the instantaneous profit can be written as a function of markups alone:

$$\Pi_t(z) = c_t(z) \Big(p_t(z) - A_t(z)E \Big) = K \mu_t(z)^{-\eta} \Big(\mu_t(z) - 1 \Big), \tag{22}$$

where K is a constant in steady state. Define the markup gap $\mu_t(z) \equiv \log(\mu_t(z)/\mu^*)$ as the log deviation of the current markup to the unconstrained markup $\mu^* \equiv \frac{\eta}{\eta - 1}$. Then a second order approximation of profits produces a quadratic form in the markup gap as follows:[14]

$$\Pi(\mu_t(z)) = C - B\mu_t(z)^2. \tag{23}$$

[14] C and B are constants. C does not affect the decisions of the firm and it is omitted for the calculations of decision rules; B captures the curvature of the profit function.

Productivity, markup estimates and uncertainty Firm z's log productivity $a_t(z) \equiv \ln A_t(z)$ evolves according to a jump-diffusion process which is idiosyncratic and independent across z:

$$da_t(z) = \sigma_f W_t(z) + \sigma_u u_t(z) dQ_t(z) \tag{24}$$

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 $u_t(z) \sim \mathcal{N}(0,1)$. Firms do not observe their productivity directly, but they have two sources of information: noisy signals $s_t(z)$ and the Poisson counter $Q_t(z)$. Signals evolve as

$$ds_t(z) = a_t(z)dt + \gamma dZ_t(z) \tag{25}$$

where $Z_t(z)$ is an independent Wiener process.[15] Firms are all ex-ante identical as they face the same parameters $\{\sigma_f, \sigma_u, \lambda, \gamma\}$, but become different ex-post as they receive different realizations of productivity and signals. From the definition of the markup gap, we have that

$$\mu_t(z) = \log p_t(z) - a_t(z) - \log E - \log \mu^*. \tag{26}$$

Notice that during inaction, the markup gap is driven by the productivity process: $d\mu_t(z) = -da_t(z)$. When the price adjusts, the markup process is reset and then it follows productivity again. By symmetry of the stochastic processes, we have that $da_t(z) = -da_t(z)$. As markup gaps and productivity mirror each other, we prefer to work with markup gaps as it facilitates the solution.

Given the assumptions, a firm's state in this economy consists of a markup gap estimate $\hat{\mu}(z)$ and its uncertainty $\Omega(z)$, where their evolution is given by the filtering equations derived in Proposition 1. Each process is indexed by z and is *iid* across firms:

$$d\hat{\mu}_t(z) = \Omega_t(z) d\hat{Z}_t(z), \qquad \hat{Z}_t(z) \sim Wiener \tag{27}$$

$$d\Omega_t(z) = \frac{\sigma_f^2 - \Omega_t^2(z)}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t(z). \tag{28}$$

Steady state equilibrium Given exogenous stochastic processes for idiosyncratic productivity $W_t(z), Q_t(z)$, idiosyncratic noise $Z_t(z)$, an equilibrium is defined by 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 $p_t(z)$, markup gap estimates $\hat{\mu}_t(z)$ and uncertainty $\Omega_t(z)$, (iii) prices P_t, E_t, R_t, D_t, and (iv) a fixed distribution over firm states $F(\hat{\mu}, \Omega)$ such that the household and firms optimize, markets clear at each date, and the distribution is consistent with actions. In a steady state equilibrium with constant money supply, the price index, nominal wages, and nominal interest rates are constant, as there is no aggregate uncertainty. Nominal expenditure is constant and equal to the nominal wage, and by market clearing, aggregate output equals aggregate consumption and the real wage $Y = \frac{E}{P}$.

[15] This signal extraction problem can be reinterpreted, in discrete time, as a problem with undistinguishable permanent and transitory shocks. The signal noise can be reinterpreted as transitory volatility. This is a useful alternative for the interpretation of the model. See Online Appendix Section G for details.

3.2 Aggregate price and output deviations

We study the real effects of a monetary shock through the following experiment. Starting from a zero inflation steady state at $t = 0$, we introduce a one-time unanticipated permanent increase in money supply of size $\delta \approx 0$, such that $\log M_t = \log \bar{M} + \delta,\ t \geq 0$. Since nominal wages are proportional to the money supply, this shock directly translates into a nominal wage increase of the same magnitude. The equality between aggregate output and the real wage implies that the increase in the nominal wage due to the monetary shock has to be distributed between price and output deviations from steady state, i.e. $\tilde{Y}_t + \tilde{P}_t = \delta$ where $\tilde{Y}_t \equiv \ln\left(\frac{Y_t}{\bar{P}}\right)$ and $\tilde{P}_t \equiv \ln\left(\frac{P_t}{\bar{P}}\right)$. Furthermore, the output gap can be approximated to a first order by:

$$\tilde{Y}_t = -\int_0^1 \mu_t(z) dz. \tag{29}$$

Finally, adding and subtracting $\hat{\mu}_t(z)$, and defining individual forecast errors as $\hat{\varphi}_t(z) \equiv \hat{\mu}_t(z) - \mu_t(z)$, we obtain the following expression:

$$\tilde{Y}_t = \underbrace{-\int_0^1 \hat{\mu}_t(z)\, dz}_{\text{average inaction error } \mathcal{I}_t} + \underbrace{\int_0^1 \hat{\varphi}_t(z)\, dz}_{\text{average forecast error } \mathcal{F}_t} = \mathcal{I}_t + \mathcal{F}_t. \tag{30}$$

Expression (30) states that output at time t differs from its steady state value if there are average pricing mistakes. Pricing mistakes arise either from "inaction errors" due to the menu cost (first term) or from forecast errors due to information frictions (second term). We call inaction errors those pricing mistakes that arise by being inside the inaction region—the firm is aware of these mistakes, which are optimal. Importantly, the output effect considers the cross-sectional average of these errors and any idiosyncratic error is washed out. Total output effects of the monetary shock are measured as the area under the impulse-response function denoted by $\mathcal{M}$ and given by

$$\mathcal{M} \equiv \int_0^\infty \tilde{Y}_t\, dt = \int_0^\infty \mathcal{I}_t\, dt + \int_0^\infty \mathcal{F}_t\, dt = \mathcal{I} + \mathcal{F}. \tag{31}$$

where we have substituted (30) and defined $\mathcal{I}$ and $\mathcal{F}$ as the present value of inaction and forecast errors, respectively. This measure is adequate as it accounts for both short- and long-run responses to the shock.[16] Note that by exchanging the order of integration between time and states in (31), we can express the total output effect as the average of the stream of individual mistakes:

$$\mathcal{M} = \int_0^1 \underbrace{\left[\int_0^\infty \hat{\varphi}_t(z)\ - \hat{\mu}_t(z)\ dt\right]}_{\text{stream of pricing mistakes by firm } z} dz. \tag{32}$$

This expression is very convenient because the stream of individual mistakes—forecast errors and inaction errors—can be characterized recursively as we show below.

[16] In the transition towards the new steady state, there are general equilibrium effects arising from changes in the average markup which affect individual policies. Proposition 7 in Álvarez and Lippi (2014) demonstrate that in this type of framework without complementarities, such general equilibrium effects can be ignored. Following their result, we compute output responses using steady state policies.

Without any frictions, firms immediately increase their price to reflect the monetary shock. There are no inaction errors at the firm-level or forecast errors cancel in the aggregate; thus the monetary shock has no real output effects. Without the menu cost, the firms' decisions are static and their markups are equal to the frictionless markup every period. In this case, as shown in Hellwig and Venkateswaran (2014), there will be no output effects from a monetary shock even in the presence of information frictions.[17] With menu costs and heterogenous uncertainty, the price level will fail to fully reflect the monetary shock and there will be real effects. We now analytically characterize these effects in three cases: a perfectly observed monetary shock, a partially observed monetary shock, and a monetary shock that is accompanied by an aggregate uncertainty shock. Each case highlights a different amplification mechanism generated by the uncertainty cycles.

The first exercise assumes that the monetary shock δ is fully observed; thus we say that it is *disclosed.* All markup estimates are fully updated and they fall by δ on impact: $\hat{\mu}_0(z) = \hat{\mu}_{-1}(z) - \delta$. Forecast errors $\hat{\varphi}_t(z)$ are *iid* across firms and the average forecast error $\mathcal{F}$ is equal to zero. Average inaction errors $\mathcal{I}$ are the only source of real output effects. Proposition 9 analytically characterizes the output response. The strategy follows Álvarez, Le Bihan and Lippi (2014) to express recursively the stream of individual pricing mistakes and aggregate them to obtain the total effect in (32).

Proposition 9 **(Real effects from a disclosed monetary shock).** *Assume the economy is in steady state and it is hit with a one-time unanticipated monetary shock of size $\delta > 0$. If firms fully observe the monetary shock, then:*

1. *The total output response consists exclusively of inaction errors, which evolve as:*

$$\textit{(total effect)} \qquad \mathcal{M}(\delta) = -\int_{\hat{\mu},\Omega} w(\hat{\mu}-\delta,\Omega)\; dF(\hat{\mu},\Omega) \tag{33}$$

$$\textit{(inaction errors by firm)} \qquad w(\hat{\mu},\Omega) = \mathbb{E}\left[\int_0^{\tau} \hat{\mu}_t \; dt \Big| (\hat{\mu}_0,\Omega_0) = (\hat{\mu},\Omega)\right] \tag{34}$$

   subject to the dynamics of $\hat{\mu}_t$ in (5) and Ω_t in (6).

2. *When $\delta > 0$, up to first order, the total output effect with uncertainty cycles is bounded below by the total effect without them ($\lambda = 0$):*

$$\mathcal{M}(\delta) \geq \delta \underbrace{\left(\frac{\mathbb{E}[\tau]}{6}\right)}_{\textit{without uncertainty cycles}}. \tag{35}$$

For computing the output effects of a disclosed monetary shock, what matters is the *first* price change of each firm, as that price change fully incorporates the monetary shock; after that, price changes respond solely to idiosyncratic conditions that wash out in the aggregate. All this is reflected in Equations (33) and (34), which only accumulate inaction errors up to the first price change by each firm. Equation (35) tells us that the output effects are bounded below by those obtained in the

[17] See Online Appendix Section G.

case without heterogeneity. This lower bound, which is proportional to the average expected time to adjustment, is a special case of the formula derived in Álvarez, Le Bihan and Lippi (2014) that relates output effects with kurtosis and frequency of the price change distribution. With these results, we now discuss two amplification mechanisms generated by the uncertainty cycles.

Amplification through dispersion in adjustment frequency. The first amplification force arises from dispersion of times until the first adjustment, and it is well-known in the literature. Output effects relate to the average expected time, that equals the average of the inverse of adjustment frequencies (not the inverse of the average frequency). By Jensen's inequality, dispersion in frequencies increases the average expected time, and in turn, amplifies the output effect. In our model, the uncertainty cycles are responsible for introducing dispersion in frequencies: high uncertainty firms have a high adjustment frequencies and almost immediately react to the money shock, while low uncertainty firms have low frequency and thus accumulate a bunch of pricing mistakes before adjusting due to their inaction.

Despite providing amplification due to dispersion in frequencies, the uncertainty cycles also introduce a dampening force due to their effect on the size of price changes: the first price change of high uncertainty firms is quite large, due to their wide inaction regions, and produces an overreaction of the aggregate price level on impact. This reminds us of the selection effect highlighted in Golosov and Lucas (2007), in which prices changes occur in the firms with the largest needs to adjust, but here belief heterogeneity drives the selection. Still, this selection effect does not undo the amplification due to dispersion in frequencies, as confirmed by the lower bound.

Amplification through a positive correlation between the strength of selection effects and adjustment frequency. The second amplification force is more subtle. The uncertainty cycles basically generate two sets of firms. The first group consists of low frequency adjusters that change their prices *primarily* due to the arrival of infrequent shocks as in Gertler and Leahy (2008); their pricing behavior features small selection effects and are responsible for most of the real effects of nominal shocks. The second group consists of high frequency adjusters that *primarily* change their prices due to the diffusion process as in Golosov and Lucas (2007); their pricing behavior features strong selection effects that dampen the real effects of nominal shocks. This means that uncertainty cycles generate endogenously a positive correlation between the strength of the selection effects and the frequency of adjustment: the low frequency adjusters are also those that does not respond much to monetary shocks. This correlation, which is a new amplification mechanism in the literature, further amplifies the real effects of monetary shocks.[18]

3.3 Partially disclosed monetary shock

The second exercise assumes that the monetary shock is only *partially disclosed.* This assumption allows us to gauge the importance of aggregate forecast errors in the propagation of the monetary shock (these were absent with a fully disclosed shock). Micro-founding this assumption is outside the

[18]See Online Appendix Section H for details.

scope of this paper, but there is an abundant literature that provides a plethora of alternatives to think about imperfect knowledge about aggregate shocks. A few examples are sticky information in Mankiw and Reis (2002), rational inattention in Woodford (2009) and Maćkowiak and Wiederholt (2009), dispersed knowledge in Hellwig and Venkateswaran (2014), level-k reasoning in Farhi and Werning (2017), among many others. In the quantitative section, we explain how we bring discipline to the observability of the monetary shock using evidence from survey forecast data.

We assume that firms only observe a fraction $\alpha \in [0,1]$ of the monetary shock δ, and that they filter the monetary shock with the same learning technology they use to estimate their idiosyncratic state. These assumptions imply the following results (see proof of Proposition 10 for details). Markup estimates are only partially updated $\hat{\mu}_0(z) = \hat{\mu}_{-1}(z) - \alpha\delta$; as markups are overestimated, positive forecast errors arise on impact $\hat{\varphi}_0(z) = (1-\alpha)\delta$ but then decrease with learning. As idiosyncratic shocks are realized, further forecast errors arise but these are unrelated to the money shock. Thus we decompose the forecast error into an unbiased and a biased component. The unbiased component is *iid* across firms and can be ignored for aggregate purposes. The biased component, denoted by φ_t follows a deterministic path for each firm: its initial value is equal across firms and then it disappears over time, at a rate that depends on idiosyncratic uncertainty as follows

$$\varphi_t(z) = \varphi_0(z) e^{-\int_0^s \frac{\Omega_t(z)}{\gamma} ds}, \quad \text{with initial bias} \quad \varphi_0(z) \sim (1-\alpha)\delta \quad \forall z. \tag{36}$$

In this scenario, average forecast errors $\mathcal{F}_t$ (defined in (30)) are no longer equal to zero. Proposition 10 establishes that average forecast errors amplify the output response—even in the absence of uncertainty cycles—but the response is amplified even more when these are present.

Proposition 10 (Real effects from undisclosed monetary shock). *Assume the economy is in steady state and it is hit with one-time unanticipated monetary shock of size $\delta > 0$. If firms only observe a fraction $\alpha \in [0,1]$ of the monetary shock, then:*

1. *The total output effect is the sum of inaction and forecast errors, which evolve as:*

$$\begin{aligned}
&(\textit{total effect}) & \mathcal{M}(\delta,\alpha) &= \mathcal{I}(\delta,\alpha) + (1-\alpha)\delta\mathcal{F} & (37)\\
&(\textit{inaction error}) & \mathcal{I}(\delta,\alpha) &= -\int_{\hat{\mu},\Omega} w(\hat{\mu}-\alpha\delta, \Omega, (1-\alpha)\delta)\; dF(\hat{\mu},\Omega) & (38)\\
& & w(\hat{\mu},\Omega,\varphi) &= \mathbb{E}\left[\int_0^\tau \hat{\mu}_t \; dt + w\left(0, \Omega_\tau, \varphi e^{-\int_0^\tau \frac{\Omega_s}{\gamma} ds}\right) \Big| (\hat{\mu}_0,\Omega_0) = (\hat{\mu},\Omega)\right] & (39)\\
&(\textit{forecast error}) & \mathcal{F} &= \int_\Omega \mathbb{E}\left[\int_0^\infty e^{-\int_0^t \frac{\Omega_s}{\gamma} ds} dt \Big| \Omega_0 = \Omega\right] dh(\Omega) & (40)
\end{aligned}$$

where Ω_t evolves as in (6) and $\hat{\mu}_t$ now follows a biased process $d\hat{\mu}_t(z) = \Omega_t(z)\left[\frac{\varphi_t(z)}{\gamma} dt + d\hat{Z}_t\right]$.

2. *Let $\underline{\mathcal{F}} \equiv \sqrt{\frac{\gamma^2 \mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}$ be a function of price statistics. When $\delta > 0$, up to a first order, the total*

output effect with uncertainty cycles is bounded below by the total effect without them ($\lambda = 0$)*:*

$$\mathcal{M}(\delta, \alpha) \geq \delta \underbrace{\left(\alpha \frac{\mathbb{E}[\tau]}{6} + (1-\alpha)\underline{\mathcal{F}} \right)}_{\text{without uncertainty cycles}}. \tag{41}$$

With a partially disclosed monetary shock, pricing mistakes do not disappear after a firm's first price change. This is evident in the recursive nature of equation (39). The monetary shock is only partially incorporated into prices with each adjustment, i.e. the passthrough of the monetary shock into prices is incomplete.

Amplification through dispersion in learning dynamics. Let us focus on the role of forecast errors. Equation (40) computes the contribution of average forecast errors to the real effects of a monetary shock. With homogenous uncertainty ($\Omega_t = \sigma_f$), average forecast errors $\mathcal{F}_t$ converge back to zero at an exponential rate $e^{-\sigma_f/\gamma}$. With heterogeneous uncertainty, there is dispersion in convergence rates: forecast errors of high uncertainty firms disappear at a faster rate than those of low uncertainty firms. This dispersion, together with Jensen's inequality, imply that the convergence of average forecast errors is slower than in the homogenous case. This slower convergence increases total output effects, as average forecast errors persist for a longer period. The value $\underline{\mathcal{F}}$ in Equation (41) provides a lower bound for forecast errors, and it can be disciplined with price statistics.[19]

An immediate consequence of the endogenous variation in the speed of firm learning is that passthrough of monetary shocks into prices is heterogenous. When firms are more uncertain, they learn more quickly about the monetary shock, increasing the responsiveness of their prices to that shock. They also respond more aggressively to idiosyncratic shocks and increase price change dispersion. This positive relationship between price change dispersion and passthrough of nominal shocks is documented empirically by Berger and Vavra (2017). As they show, such a relationship naturally arises in models with time-varying responsiveness, as in our Bayesian learning model with uncertainty cycles.

3.4 Aggregate uncertainty shock

Finally, our third exercise explores the output effects of a monetary shock that occurs at the same time as an aggregate uncertainty shock. The analysis is motivated by the empirical finding that monetary policy is less effective when economic uncertainty is higher. We interact an undisclosed monetary shock with an unanticipated uncertainty shock that increases every firm's uncertainty by a multiple κ of average uncertainty $\mathbb{E}[\Omega]$. An example of this type of shock is a monetary expansion during a recession or any period of elevated economic uncertainty. Proposition 11 shows that larger *average* uncertainty decreases real output effects; still, uncertainty cycles maintain their role in amplifying output effects compared to an economy with homogenous uncertainty.

[19] For example, the lower bound $\underline{\mathcal{F}}$ decreases with price change dispersion $\mathbb{V}[\Delta p]$, i.e. larger price change dispersion reduces the amplification potential of forecast errors.

Proposition 11 (**Real effects from monetary and uncertainty shock**)**.** *Assume the economy is in steady state and it is hit with one-time unanticipated monetary shock of size $\delta > 0$. Simultaneously, idiosyncratic uncertainty increases by $\kappa\mathbb{E}[\Omega]$ for all firms. If firms only observe a fraction $\alpha \in [0,1]$ of the monetary shock, then:*

1. *The total output effect $\mathcal{M}(\delta,\alpha,\kappa) = \mathcal{I}(\delta,\alpha,\kappa) + (1-\alpha)\delta\mathcal{F}(\kappa)$ is computed as in Proposition 10, but with an initial condition for uncertainty that reflects the uncertainty shock: $\Omega_0 = \Omega + \kappa\mathbb{E}[\Omega]$.*

2. *When $\delta > 0$, the forecast error component $\mathcal{F}(\kappa)$, now a function of κ, is bounded below by*

$$\mathcal{F}(\kappa) \geq \xi\left(\frac{\mathbb{E}[\tau]\mathbb{E}[\Omega]^2(1+\kappa)^2}{\mathbb{V}[\Delta p]}\right)\underline{\mathcal{F}} \geq \underbrace{\xi((1+\kappa)^2)\ \underline{\mathcal{F}},}_{\textit{without uncertainty cycles}} \tag{42}$$

where $\xi(\cdot)$ is decreasing, $\lim_{x\to 1}\xi(x) = 1$ and $\lim_{x\to\infty}\xi(x) = 0$.

The aggregate uncertainty shock shifts the uncertainty distribution to the right by $\kappa\mathbb{E}[\Omega]$. This shift increases the flexibility of the aggregate price level on impact by increasing the mass of high uncertainty firms and the implied selection effect. This reduces the short-run response of output. Vavra (2014) studies this effect in a menu cost model with aggregate volatility shocks. Moreover, the shift in the uncertainty distribution decreases the persistence of forecast errors by reducing the mass of low uncertainty firms. This reduces the long-run response of output. Besides the effects that arise from a shift in the distribution, there is an additional force that affects all firms' learning dynamics. Firms place a higher weight on news when their uncertainty is higher; this is what Bayesian updating is all about. Forecast errors disappear faster and the monetary shock is quickly incorporated into prices. In turn, this reduces the persistence of the output response. This learning force is evident in equation (42), the lower bound of forecast errors, which is decreasing in the size of the aggregate uncertainty shock κ. Overall, monetary shocks are less effective at increasing output when aggregate uncertainty is higher, in line with empirical research. Our results also relate to the findings in Coibion and Gorodnichenko (2015) which show that information rigidities are lower during periods of high uncertainty, just as our model predicts.

4 Quantifying the mechanisms

In this section we quantify the mechanisms responsible for the amplification of real output effects with heterogenous uncertainty. For this purpose, we solve the model and calibrate it to match micro price statistics in the US Consumer Price Index.[20]

4.1 Data and calibration

The calibration is at the weekly frequency. The discount factor is set to $\frac{1}{1+r} = 0.96^{1/52}$ to match an annual risk free rate of 4%; the CES elasticity is set to $\eta = 6$ to match an average markup of 20%;

[20]The model is solved numerically in discrete time. See Section I in the Online Appendix for details.

and the disutility of labor is set to $\alpha = 1$. Following the empirical evidence, we set the normalized menu cost to $\bar{\theta} = 0.064$ so that the expected menu cost payments represent 0.5% of average revenue; price statistics and impulse-responses are robust to alternative values for this parameter.[21]

Price statistics computed in the model (at weekly frequency) are aggregated to match the monthly price statistics in the data. We target statistics in the BLS data computed in Nakamura and Steinsson (2008): price change dispersion $std[|\Delta p|] = 0.08$, average price duration of $\mathbb{E}[\tau]$=10 months, and hazard rate's average slope between 1 and 18 months $slope = -0.007$. We consider three parametrizations of the model, of all which match the same average price duration. Table I summarizes the data, the target moments in each calibration (marked with a star), and other non-targeted moments produced.

Table I – Parameters and Targets

	US Data	(1) Baseline	(2) Fat-tailed shocks	(3) Heterogenous uncertainty
Parameters				
σ_f		0.016		0
σ_u			0.146	0.198
λ			0.035	0.016
γ				0.233
Moments				
$\mathbb{E}[\tau]$ in months	10	10^*	10^*	10^*
std$[\|\Delta p\|]$	0.08	0.007	0.08^*	0.07^*
hazard rate slope	-0.007	0.007	0.000	-0.009^*
kurtosis$[\Delta p]$	3.95	1.027	2.26	1.96

Data: US CPI data. Models: (1) Baseline: Perfect info and frequent shocks; (2) Perfect info and fat-tailed shocks; (3) Heterogeneous uncertainty. Hazard rate's slope - average between 1 and 18 months. * = targeted moment.

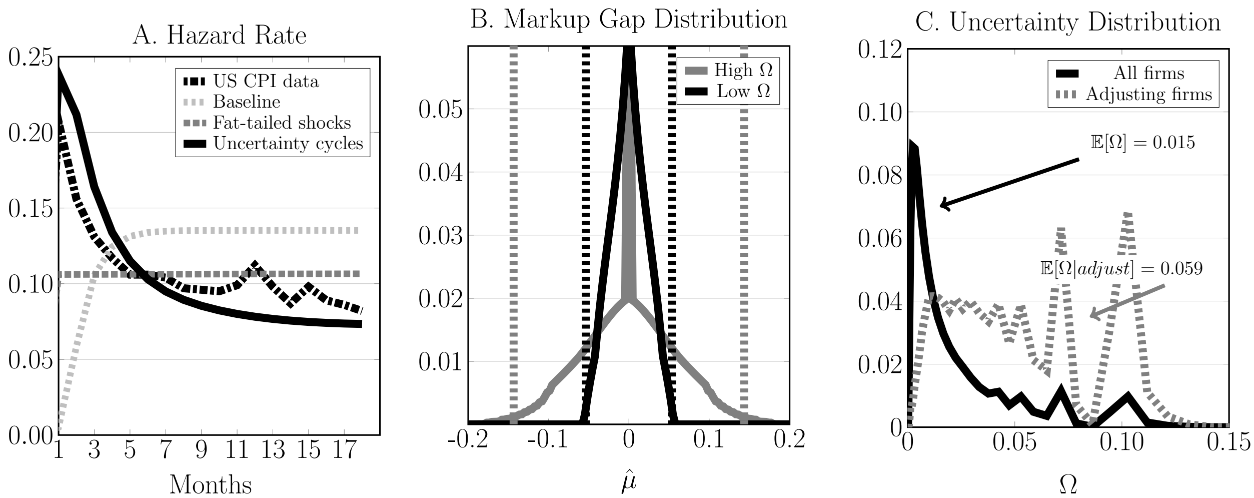
The baseline calibration in Column (1) shuts down the information friction ($\gamma = 0$) and fat-tailed shocks ($\lambda = 0$). Its only parameter σ_f is set to match average duration. We consider this as a simplified version of Golosov and Lucas (2007). Column (2) shuts down the information friction ($\gamma = 0$) and the frequent shocks ($\sigma_f = 0$), and only has fat-tailed shocks. This is a simple version of Gertler and Leahy (2008). Its two parameters, λ and σ_u, are set to match average duration and price change dispersion. Column (3) gives the numbers for our model with heterogenous uncertainty. We calibrate it using a simulated method of moments that matches average duration, price change dispersion, and the hazard rate's slope. Importantly, we match the same duration with an arrival rate λ that is 1/3 of the arrival rate in the model without heterogeneity. For each fat-tailed shock, prices change more than once because of the decreasing hazard; this is key to amplify the persistence of the output response. Finally, the calibration sets the volatility of the frequent shocks, σ_f, very close to zero.

Figure III shows the hazard rate and the steady state distributions. Panel A plots the hazard rate for the US data and the three parameterizations. The baseline calibration features an increasing hazard rate, the calibration with only fat-tailed shocks produces a flat hazard, and the calibration with heterogenous uncertainty generates the non-monotonic hazard with the decreasing segment, as

[21] Levy *et al.* (1997) estimates that the cost of changing prices is about 0.7% of firms' revenue for supermarket chains. For a large retailer, Zbaracki *et al.* (2004) estimates 0.04% for physical costs, 0.28% for managerial costs, and 0.89% for customer costs. We choose a number in between these estimates as our baseline, as in Golosov and Lucas (2007) and Gertler and Leahy (2008), and provide a robustness analysis in Section J.1 of the Online Appendix.

in the data. Panel B plots the distribution of markup gap estimates for high and low uncertainty firms. The distribution's support and dispersion increase with uncertainty. Average inaction regions are $|\bar{\mu}(\Omega)| = 0.14$ and $|\bar{\mu}(\Omega)| = 0.06$ for high and low uncertainty firms, respectively. Panel C shows the two uncertainty distributions. Consistent with Proposition 7, the steady state distribution of uncertainty $h(\Omega)$ is biased towards low uncertainty with an average level of 0.015. In contrast, the renewal distribution $r(\Omega)$ shifts its mass towards higher levels, with an average uncertainty of 0.059.[22]

Figure III – Hazard Rates and Steady State Distribution



Data: US CPI. Models: High uncertainty > 80th percentile, Low uncertainty < 20th percentile.

Robustness with respect to the adjustment hazard. Our calibration targets the adjustment hazard in Nakamura and Steinsson (2008). Section J.2 in the Online Appendix analyzes quantitatively the role of uncertainty cycles for different calibrations that target other empirical hazards. We show that our mechanism is still present, although it gets amplified or dampened, when targeting a steeper hazard as in Campbell and Eden (2014), or a flatter hazard as in Klenow and Kryvtsov (2008). An interesting feature is that a highly-decreasing hazard is matched with a relatively low signal noise γ, because with little noise it becomes easier to learn the markup, and both the uncertainty and the adjustment probability fall quickly. It follows that matching a steep hazard rate generates larger heterogeneity in uncertainty, more dispersion is adjustment frequency, larger kurtosis for the price change distribution, and a stronger correlation between selection effects and adjustment frequency. All of these channels together amplify the output effects of a monetary shock, which get closer to the case with a constant hazard (Calvo).

[22]The model with heterogenous uncertainty generates larger kurtosis than the baseline but below the kurtosis in the data. Section K in the Online Appendix considers an extension with random opportunities to freely adjust. This extension increases the kurtosis by generating more small price changes. 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).

4.2 Further inspection of information frictions in the data

Before the quantitative exercise, we perform a deeper inspection of the micro data through the lens of our model with the objective of providing further evidence on the presence of information frictions. We discover a connection between price "age" and uncertainty; we measure the profit losses due to the information friction; and we explain how to put discipline on the observability of the monetary shocks using survey data.

Age-dependent statistics. As a robustness check of the calibration, we exploit the implication 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 (recently set) and old prices (set many periods ago) exhibit different behavior. In particular, young prices are more dispersed and more likely to be reset than old prices. These predictions are documented by Campbell and Eden (2014) using weekly scanner data. They find 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. They also find 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 20th quintile of the price age distribution and old prices to be in the 80th quintile. We obtain that price dispersion is 1.3 times higher and adjustment frequency is 2 times higher for young prices. Interestingly, the average uncertainty associated to young prices is 3.3 times the uncertainty faced by old prices, thus relative price dispersion and adjustment frequency are informative about relative uncertainty.[23]

Losses due to information friction. Using the model and micro-price statistics, we can gauge the size of the profit losses that are due to information frictions and nominal rigidity.[24] Proposition 12 characterizes these losses, and provides an upper bound in terms of cross-sectional moments and structural parameters.

Proposition 12 (**Revenue losses due to frictions**). *Consider a constant returns to scale technology and a CES demand with elasticity $\eta > 1$. Then we can express the expected per period profit losses that arise from frictions relative to the frictionless benchmark, expressed as a fraction of revenue, as:*

$$Losses = \frac{1-\eta}{2}\left[\underbrace{\gamma\mathbb{E}\left[\Omega\right]}_{\textit{Loss from info. friction}} + \underbrace{\mathbb{V}\left[\hat{\mu}\right]}_{\textit{Loss from nominal rigidity}}\right] \tag{43}$$

Moreover, the losses due to the information friction are bounded above by $\frac{1-\eta}{2}\gamma\sqrt{\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}}$.

The first term shows the loss that arises from ignoring the true state, while the second term shows the loss that arises from setting an incorrect price, reflected through markup dispersion. In our preferred calibration, the term related to information frictions equals 0.86% of revenues, with an

[23] See Online Appendix Section L for details on the age-uncertainty relationship and age dependent statistics.

[24] We thank one of our referees for suggesting this analysis.

upper bound of 1.6% of revenue.[25] While these numbers appear to be slightly higher compared to the 0.28% of managerial costs in Zbaracki *et al.* (2004), which include some information-gathering costs, we think that further empirical evidence is needed in order to pin down this number with better precision.

Discipline on forecast errors and observability of the monetary shock. Through the model, the micro-pricing data tightly disciplines the forecast error dynamics. Our estimates for γ and the implied uncertainty distribution suggest forecast error persistence at the individual level of 0.94 for the average firm and 0.84 for the average high uncertainty firm.[26]. The persistence of aggregate forecast errors is 0.9. These numbers are consistent with the evidence in Coibion and Gorodnichenko (2012), who estimate forecast error persistence in the range of 0.75 to 0.9 using survey data from US professional forecasters, consumers, firms, and central bankers. Furthermore, their results suggest that Bayesian forecasters assign a weight lower than 0.2 on new information. Our calibration is also consistent with this, as we obtain a weight on new information between 0.06 (for an average firm) and 0.16 (for a high uncertainty firm). The weight on new information implies that it takes about three quarters to reduce the forecast error by a half, just as our quantitative exercise suggests (Panel C in Figures V and VI).

The convergence of forecast errors is independent from the observability of the monetary shock α; nevertheless, observability determines the initial level of the average forecast error following a monetary shock. Again, we use the survey evidence in Coibion and Gorodnichenko (2012) to discipline this parameter. They find that, on average, the initial response of forecast errors is about half the response of the forecasted variable. Given this evidence, we set observability equal to $\alpha = 0.5$. The observability implied by survey data is relatively high when compared to the rational inattention literature, which suggests that aggregate shocks are almost undisclosed, e.g. Maćkowiak and Wiederholt (2009)'s model implies that $\alpha = 0.04$. For this reason, we additionally show results for a fully undisclosed shock ($\alpha = 0$), so that the whole spectrum of observability is covered in our quantitative analysis.

4.3 Quantifying the real effects of monetary shocks

We proceed to quantify the real effect of monetary shocks in the three cases we theoretically analyzed: disclosed monetary shocks, partially disclosed monetary shocks, and monetary shocks accompanied by an aggregate uncertainty shock.

The role of selection effect. We start by quantifying the output effect of a one-time unanticipated increase in money supply of size $\delta = 1\%$ that is fully disclosed by the monetary authority. Figure IV shows the impulse-response of output for the three calibrations.[27] In the baseline with perfect

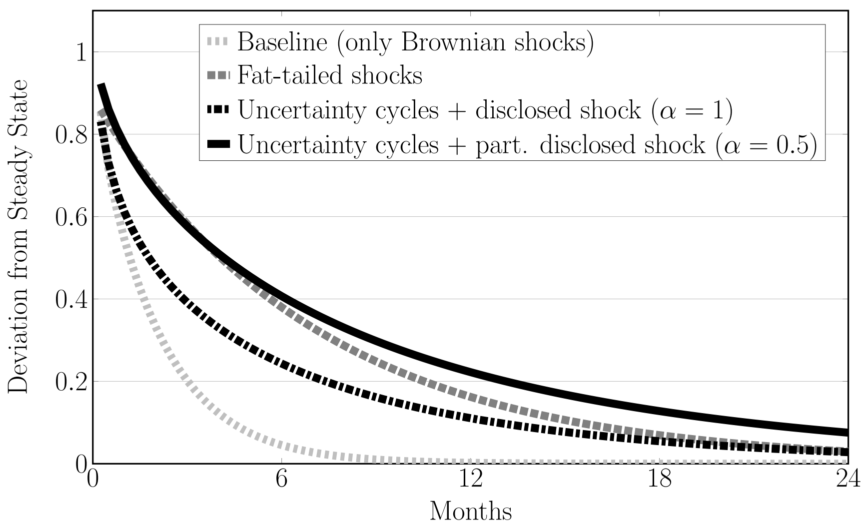
[25]The numbers are computed as: $\frac{1-\eta}{2}\gamma\mathbb{E}\left[\Omega_t(z)\right] = -\left(\frac{1-6}{2}\right)*0.233*0.0147 = -0.0086$, and the upper bound is computed with weekly price statistic as $\frac{1-\eta}{2}\gamma\sqrt{\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}} = -\left(\frac{1-6}{2}\right)*0.233*\sqrt{\frac{0.03}{40}} = -0.016$.

[26]Forecast error persistence is equal to the Bayesian weight assigned to prior information: $\frac{\gamma^2}{\gamma^2+\Omega_t(z)^2}$. We compute the cross-sectional average of this number for all firms and for firms with uncertainty above the mean.

[27]Note that all the impulse-responses have a jump on impact. While this is not surprising for the second calibration with only fat-tailed shocks (there is a positive mass of firms at the borders of inaction), this jump does not occur in

information, the monetary shock of 1% generates a total output effect of $\mathcal{M} = 1.61\%$. This number is very close to that implied by the lower bound formula in Proposition 9, which yields: $\delta \frac{\mathbb{E}[\tau]}{6} = 1.66\%$. The small and short-lived output response (1.25 months) is the result of a large selection effect. The second calibration with fat-tailed shocks more than triples the baseline's output effects and persistence. Its flat hazard breaks the selection effect and obtains larger non-neutrality. The model with heterogenous uncertainty and fully disclosed monetary shock more than doubles the output effects in the baseline, but the half-life is only 40% larger. The impulse-response almost perfectly tracks the baseline in the first months, then it slowly approaches the impulse-response of the second model, and eventually crosses it at 24 months. The quick and large adjustments of high uncertainty firms dominate and drastically reduce the output effect during the first months; thus the short half-life. Nevertheless, the total output effect is amplified by heterogeneity as demonstrated in Proposition 9, thanks to the slow responsiveness of low uncertainty firms.[28] Lastly, the model with heterogeneous uncertainty and a partially disclosed shock, as calibrated with the survey data, obtains real output effects that surpass the other models.

Figure IV – Output Response to a Monetary Shock



(1) Baseline: perfect info with only frequent shocks; (2) Perfect info with only fat-tailed shocks; (3) Heterog. uncertainty + disclosed shock; (4) Heterog. uncertainty + partially disclosed shock (from survey data).

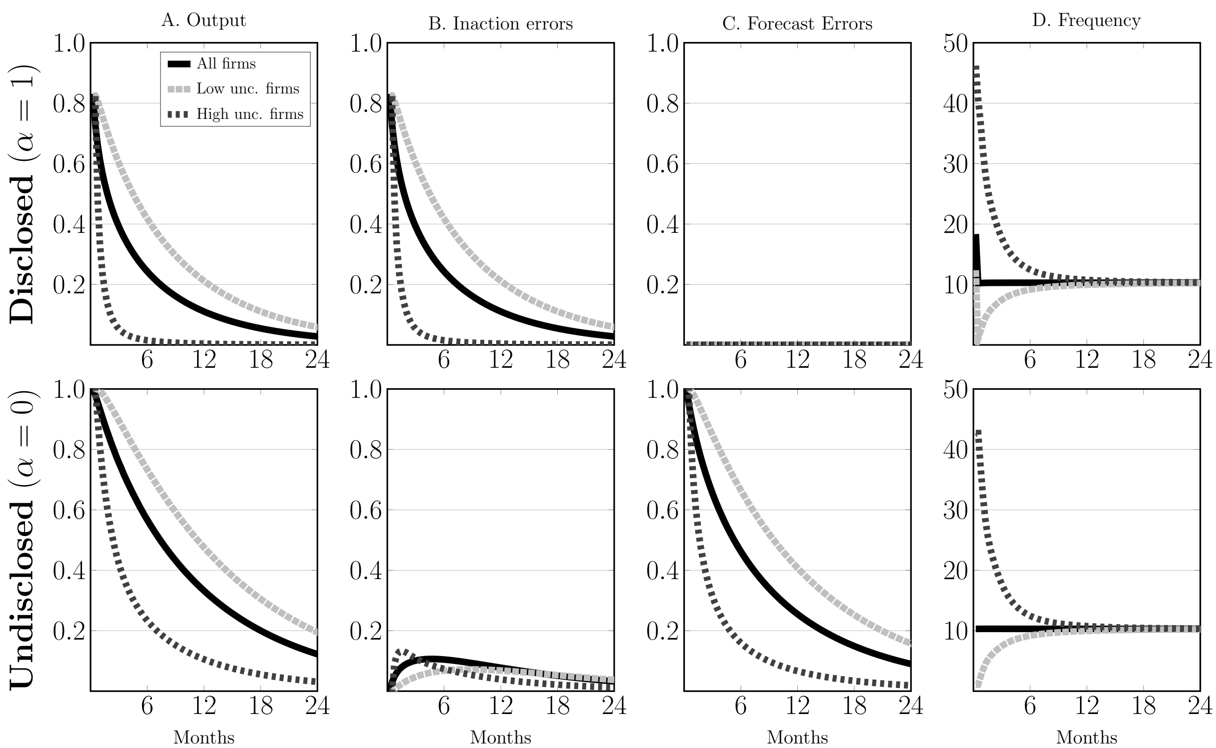
The role of forecast errors. To highlight the role of average forecast errors, we measure the output effect of a one-time unanticipated increase in money supply of size $\delta = 1\%$ that is fully undisclosed by the monetary authority ($\alpha = 0$). The lack of observability is in line with the rational inattention

the other models with Brownian shocks as the impact of a small monetary shock is second order. Here, the jumps arise because we solve the model in discrete time and there is a positive mass of firms at the borders of inaction in all calibrations. See Section M in the Online Appendix for more details on the computation of IRFs and a detailed comparison of the impact responses across models.

[28] Note that, quantitatively, a large fraction of the amplification is due to the leptokurtic shocks. Since Gertler and Leahy (2008) and Midrigan (2011), it is well known that a menu cost model with fat-tailed shocks behaves very closely to a Calvo model, where selection is minimal and output effects are large. Our results emphasize that the idiosyncratic uncertainty cycles introduce amplification above and beyond what is already embedded in a particular model.

literature. In this case, output effects are mainly driven by forecast errors. Figure V plots output deviations from steady state for the extreme cases of fully disclosed shock ($\alpha = 1$) and undisclosed shock ($\alpha = 0$), and decomposes the deviations into inaction errors and forecast errors. We separately display the aggregate response, the response of high uncertainty firms, and the response of low uncertainty firms, where averages are conditional on uncertainty at the moment of the monetary shock.

Figure V – Output Response Conditional on Uncertainty and Observability of Monetary Shock

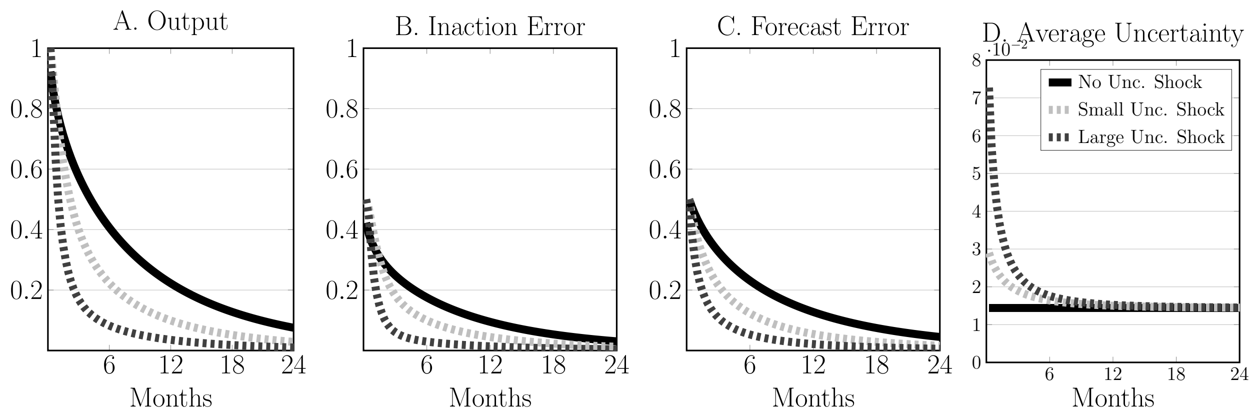


Impulse-responses after a monetary shock. The first three variables are measured as deviations from steady state, while adjustment frequency is plotted in levels. Column A is equal to the sum of Columns B and C. First row = disclosed, second row = undisclosed. Responses are conditional on initial uncertainty: solid line = total mass of firms, light dashed line = uncertainty below 20th percentile, dark dotted dashed line = uncertainty above 80th percentile.

Comparing across rows, we observe that the output response and half-lives are evidently larger with an undisclosed shock. Note that forecast errors of highly uncertain firms quickly decrease; these firms learn and incorporate quickly the monetary shock. Low uncertainty firms drive most of the persistence in average forecast errors as they take a lot of time to learn the monetary shock; these firms are pretty certain about their own conditions and make little use of new information. Again, this is all pure Bayesian updating. The heterogeneity in learning dynamics and in passthrough amplifies the output response. According to our formula in (41), the forecast errors have to be at least as large as those that arise without heterogeneity: $\underline{\mathcal{F}} \equiv \frac{\gamma\sqrt{\mathbb{E}[\tau]}}{std[\Delta p]} = \frac{0.233\sqrt{10}}{0.1456} = 5.06$, i.e. output effects must be 5 times as large as the monetary shock. Numerically, we obtain an output response 7 times larger than the monetary shock. We interpret these numbers as saying that heterogeneous uncertainty amplifies the output response in 40%.

Simultaneous monetary and uncertainty shock. Finally, we compute the effect of a partially disclosed monetary shock ($\alpha = 0.5$) that occurs at the same time of an aggregate uncertainty shock. We assume that the uncertainty shock increases firms' uncertainty by $\kappa\mathbb{E}[\Omega]$, where $\mathbb{E}[\Omega]$ is average steady state uncertainty and $\kappa \in \{0, 1, 4\}$. Figure VI shows the results.

Figure VI – Output Response to Simultaneous Monetary and Uncertainty Shock



Impulse-response after a partially disclosed monetary shock ($\alpha = 0.5$) paired with an aggregate uncertainty shock. The first three variables are measured as deviations from steady state, while uncertainty is plotted in levels.

Aggregate uncertainty shocks significantly reduce the output response through their effect on forecast errors, as these converge faster to zero due to the Bayesian learning forces. According to Proposition 11, the lower bound on forecast errors (which is achieved without heterogeneity) decreases with the size of the aggregate uncertainty shock κ. Thus the potential of forecast errors to amplify the output response are lower with higher aggregate uncertainty. Consistent with the theory, in our simulation we obtain that output effects fall from 4.64 times the baseline, to 2.75 times the baseline with a small uncertainty shock, and to only 1.3 times the baseline with a large uncertainty shock. The panel in the far right shows that uncertainty shocks are short-lived, as average uncertainty converges back to its steady state level after a few months.

Table II summarizes all the quantitative results. It shows the total output effect $\mathcal{M}$ and the persistence of the response measured as the half-life of the impulse-response function. Note that in this table all the numbers are multiples of the baseline case.

Table II – Output Response to Monetary Shock for Different Parametrizations

	Perfect Info		**Heterogenous uncertainty**			
	(1)	(2)	(3)	(4)	(5)	(6)
Output Effect	Baseline	Fat-tailed shocks	Disclosed money shock	Undisclosed money shock	Small Unc. shock	Large Unc. shock
	$\gamma = \lambda = 0$	$\gamma = 0$	$\alpha = 1$	$\alpha = 0.5$	$\kappa = 1$	$\kappa = 4$
Total effect	1.00	3.75	2.70	4.64	2.75	1.31
Half-life	1.00	3.40	1.60	3.40	1.80	1.00

Multiples of the baseline case in Column (1) with perfect information and only small frequent shocks. For that case, the total output effects are $\mathcal{M} = 1.61\%$ and the half-life is 1.25 months.

4.4 Intensive and Extensive Margin Decomposition

As a final note, we assess the relative importance of the intensive and extensive margins of adjustment that follow a monetary shock. The intensive margin describes the additional price increase of those firms that were going to adjust anyway; in contrast, the extensive margin, reflects the increase in the fraction of adjusters. While we cannot apply directly the first-order decomposition in Caballero and Engel (2007) because in our model a monetary shock only has second-order effects, we propose an alternative decomposition along the same lines. Table III computes these adjustment margins after a disclosed monetary shock $\delta = 1\%$ paired with an aggregate uncertainty shock of different sizes.[29]

Table III – Relative Importance on Intensive and Extensive Margins on Impact

On Impact:	No Ω shock $\kappa = 0$	Small Ω shock $\kappa = 2$	Large Ω shock $\kappa = 4$
Intensive/Total	30%	39%	44%
Extensive/Total	70%	61%	56%

We observe that, on impact, the extensive margin is larger than the intensive margin, but its relative importance is decreasing in the size of the aggregate uncertainty shock. Although an in-depth analysis of these margins is outside the scope of this paper, we think these numbers are an interesting starting point for further empirical analysis.

5 Conclusion

We have developed a new framework that combines an inaction problem, arising from a non-convex adjustment cost, together with a signal extraction problem in which agents face undistinguishable transitory and persistent shocks with jumps. 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, fat-tailed shocks, and information frictions are likely to coexist. Particularly, we foresee applications in setups that generate strong age-dependent statistics, such as labor markets. Moreover, the filtering results can be extended to aggregate shocks to study, for instance, disaster risk in a tractable way.

Firm-level uncertainty appears to be a quantitatively important determinant of the effects of monetary policy. Going forward, it would be interesting to further explore empirically the model's implications. For this purpose, it will be key to generate panel data with measures of idiosyncratic uncertainty about both aggregate and individual variables, and then pair these with pricing data and other firm characteristics. The work by Bachmann, Born, Elstner and Grimme (2017a) pushes towards this direction by combining firm-level pricing data with a qualitative measure of idiosyncratic

[29] Section N of the Online Appendix provides all the details regarding the computation of the two margins. We thank one of our referees for suggesting this exercise.

uncertainty faced by German firms. It would be interesting to incorporate such cross-sectional measures of uncertainty into VAR analyses that test for state-dependent effectiveness of monetary policy in times of high and low uncertainty, to complement the results obtained with measures of aggregate uncertainty (for an example, see Castelnuovo and Pellegrino 2018). Finally, surveys that elicit firm expectations about inflation and other aggregate variables, as in Coibion, Gorodnichenko and Kumar (2018) for the case of New Zealand, will allow constructing aggregate forecast errors to test for state dependency in learning and pricing.

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.

Alvarez, F., Le Bihan, H. and Lippi, F. (2016). The real effects of monetary shocks in sticky price models: a sufficient statistic approach. *The American Economic Review*, **106** (10), 2817–2851.

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

—, 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.

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. (2017a). Time-varying business volatility, price setting, and the real effects of monetary policy. *CEPR Discussion Paper Series*, **DP9702**.

—, Elstner, S. and Hristov, A. (2017b). Surprise, surprise-measuring firm-level investment innovations. *CEPR Discussion Paper No. DP9894*.

—, — and Sims, E. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, **5** (2), 217–249.

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

Berger, D. and Vavra, J. (2017). *Shocks vs. Responsiveness: What Drives Time-Varying Dispersion?* Tech. rep., National Bureau of Economic Research.

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

Bouakez, H., Cardia, E. and Ruge-Murcia, F. (2014). Sectoral price rigidity and aggregate dynamics. *European Economic Review*, **65**, 1–22.

Caballero, R. J. and Engel, E. M. (2007). Price stickiness in ss models: New interpretations of old results. *Journal of Monetary Economics*, **54**, 100 – 121.

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.

Castelnuovo, E. and Pellegrino, G. (2018). Uncertainty-dependent effects of monetary policy shocks: A new keynesian interpretation. *Journal of Economic Dynamics and Control*, **Forthcoming**.

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, **120** (1), pp. 116–159.

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

—, — and Kumar, S. (2018). How do firms form their expectations? new survey evidence. *American Economic Review*, **Forthcoming**.

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*.

Davis, S. J. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, **107** (3), 819–863.

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.

Farhi, E. and Werning, I. (2017). *Monetary Policy, Bounded Rationality, and Incomplete Markets*. Tech. rep., National Bureau of Economic Research.

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. (2014). Dispersed information, sticky prices and monetary business cycles: A hayekian perspective. *NYU mimeo*.

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

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.

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.

— and Malin, B. A. (2011). Microeconomic evidence on price-setting. *Handbook of Monetary Economics*, pp. 231–284.

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

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.

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.

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

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.

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.

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

Preliminaries Throughout the proofs, we denote partial derivatives with $f_{\hat{\mu}^i\Omega^j} \equiv \frac{\partial^{i+j} f}{\partial \hat{\mu}^i \partial \Omega^j}$. We denote with $\mathcal{A}$ the infinitesimal generator of $(\hat{\mu}, \Omega)$ and $\mathcal{A}^*$ its adjoint operator, which are given respectively by

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

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

See Section C of the Online Appendix for their derivation.

Proposition 1 (Filtering Equations, Including Drift). *Proposition 1 is proved in a more general setup than in the text, adding a non-zero drift for the state. Let the following processes define the state and the signal*

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

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.

(1) 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 $\mathcal{Q}_t = \sigma\{Q_r | r \le t\}$, follows a Gaussian process.

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

(3) The recursive estimation formulas obtained with discrete sampling converge to (A.4).[30]

Now we elaborate on the three steps.

Lemma 1. *Let $M_t \equiv [\mu_t, s_t]$ be the solution to (B.13) and $\mathcal{Q}_t = \sigma\{Q_r | r \le t\}$. 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's 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 induction hypothesis, assume that $M_r^k|\mathcal{Q}_t \equiv [\mu_r^k, s_r^k|\mathcal{Q}_t]$ is Normal for all $r \le t$. 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

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

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. □

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:

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

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: $\mathcal{I}_t^\infty = \sigma\{s_h, Q_h : h \in [0,t]\}$. Therefore, the estimate obtained with the limit of discrete sampling converges (in L^2) to the estimate with continuous sampling, i.e. $M_t^\infty \to_{L^2} \mu_t|\mathcal{I}_t \ \sim\ \mathcal{N}\left(\mathbb{E}\left[\mu_t|\mathcal{I}_t\right], \mathbb{E}\left[\left(\mu_t - \mathbb{E}\left[\mu_t|\mathcal{I}_t\right]\right)^2|\mathcal{I}_t\right]\right)$ (see Davis (1977) for more details in this topic). □

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)*

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

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

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

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, the previous system can be written as

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

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 P of the Online Appendix for details, where we follow Theorem 1.1, Chapter 10 of Kushner and Dupuis (2001)). $\square$

To conclude, we use the structure of the signal to rewrite the law of motion in terms of 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.

$\square$

Proof of Proposition 2 (Stopping time problem). 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:

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

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:

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

where the sum is telescopic and all terms except the first cancel out. Using the previous results, the sequential problem in (A.8) is written in terms of estimates as:

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

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. □

Proof of Proposition 3 (HJB Equation, Value Matching and Smooth Pasting). We apply Theorem 2.2 of Øksendal and Sulem (2010) to obtain sufficient conditions for optimality, which are: (1) HJB equation; (2) Value matching; and (3) Smooth pasting condition in each dimension. Using the infinitesimal 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:

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

Finally, we 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 D of the Online Appendix verifies assumptions for the sufficient conditions to hold in our problem; and Section E.3 verifies numerically that the smooth pasting conditions for $\hat{\mu}$ and Ω are valid. □

Proof of Proposition 4 (Inaction Region). 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.

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

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

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

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

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

Where we used the property of symmetry of the value function to obtain a reset markup equal to zero.

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: $V_\mu(\hat{\mu},\Omega) = V_{\hat{\mu}^2}(0,\Omega)\hat{\mu} + \frac{V_{\hat{\mu}^4}(0,\Omega)}{3!}\hat{\mu}^3$. Evaluate at the border, multiply both sides by $\frac{\bar{\mu}(\Omega)}{2}$ and combine with the smooth pasting condition (A.11) to obtain:

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

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

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

where $\Omega^{*2} \equiv \sigma_f^2 + \lambda\sigma_u^2$. Lastly, taking the limit $r \to 0$, evaluating at $\hat{\mu} = 0$ and using the property that the payoff function is bounded in the continuation region:

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

Thus

$$\bar{\mu}(\Omega) = \left(\frac{6\bar{\theta}\Omega^2}{1+\mathcal{L}^{\bar{\mu}}(\Omega)}\right)^{1/4} \quad \textit{with} \quad \mathcal{L}^{\bar{\mu}}(\Omega) \equiv \frac{\Omega^2-\Omega^{*2}}{\gamma}\frac{V_{\hat{\mu}^2\Omega}(0,\Omega)}{2} \tag{A.18}$$

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:

$$\bar{\mu}(\Omega) = (6\bar{\theta}\Omega^2)^{1/4}\left[1+\left(\frac{8}{3}\frac{\bar{\theta}}{\gamma^2}\right)^{1/2}\left(\frac{\Omega}{\Omega^*}-1\right)\right]^{-1/4} \tag{A.19}$$

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

Clearly, $\mathcal{E} < 1$. 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), $0 = V_\Omega(\bar{\mu}(\Omega), \Omega) - V_\Omega(0,\Omega)$.

□

Proof of Proposition 5 (Conditional Expected Time). Let $T(\hat{\mu},\Omega)$ denote the expected time for the next price change given the current state, i.e. $\mathbb{E}[\tau|\hat{\mu},\Omega]$. 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. 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 approx, and border condition**. The expected time between price changes satisfies

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

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:

$$0 = 1 + \frac{\Omega^{*2} - \Omega^2}{\gamma} T_\Omega(\hat{\mu},\Omega) + \frac{\Omega^2}{2} T_{\hat{\mu}^2}(\hat{\mu},\Omega) \tag{A.20}$$

The border condition states that at the border of action, the expected time is equal to zero $T(\bar{\mu}(\Omega),\Omega) = 0$.

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: $T(\hat{\mu},\Omega) = T(0,\Omega) + \frac{T_{\hat{\mu}^2}(0,\Omega)}{2}\hat{\mu}^2$. We evaluate this expression at $(\bar{\mu}(\Omega),\Omega)$ and we use the border condition to compute $T(0,\Omega)$, $T(0,\Omega) = -\frac{T_{\hat{\mu}^2}(0,\Omega)}{2}\bar{\mu}(\Omega)^2$. We evaluate the HJB in (A.20) at $(0,\Omega)$ and solve for it to compute $T_{\hat{\mu}^2}(0,\Omega)/2$, $\frac{T_{\hat{\mu}^2}(0,\Omega)}{2} = -\frac{1}{\Omega^2}\left[1 + T_\Omega(0,\Omega)\frac{\Omega^{*2}-\Omega^2}{\gamma}\right]$. Substitute both terms into the Taylor approximation and rearrange:

$$T(\hat{\mu},\Omega) = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}\left(1 + \mathcal{L}^\tau(\Omega)\right) \tag{A.21}$$

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:

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

To characterize $T_\Omega(0,\Omega^*)$, take the partial derivative of (A.21) 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.

□

Proof of Proposition 6 (Conditional Hazard Rate). 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:

$$\mathbb{E}_0[\mu_\tau^2] = \mathbb{E}_0[(\mu_0+\sigma_f W_\tau)^2] = \mathbb{E}_0[\mu_0^2+2\mu_0\sigma_f\mathbb{E}_0[(W_\tau-W_0)]+\sigma_f^2\mathbb{E}_0[(W_\tau-W_0)^2] = \mathbb{E}_0[\mu_0^2]+\sigma_f^2\tau = \Sigma_0+\sigma_f^2\tau \quad \text{(A.22)}$$

where we have use the properties of the Wiener process.

2. **Estimate's conditional variance**. The conditional forecast variance evolves as $d\Sigma_\tau = \left(\sigma_f^2 - \frac{\Sigma_\tau^2}{\gamma^2}\right)d\tau$. Assuming an initial condition Σ_0 such that $\Sigma_0 > \gamma\sigma_f$, the general solution to the differential equation is in the family of hyperbolic tangent and hyperbolic cotangent. Since the family of hyperbolic tangent does not satisfy the boundary condition, the solution is given by $\Sigma_\tau = \sigma_f\gamma\coth\left[\sigma_f\gamma c + \frac{\sigma_f}{\gamma}\tau\right]$. Evaluating at the initial condition, we get $\Sigma_0 = \sigma_f\gamma\coth[\gamma\sigma_f c]$ and therefore $c = \frac{1}{\sigma_f\gamma}\coth^{-1}\left(\frac{\Sigma_0}{\sigma_f\gamma}\right)$. Back into (2) and using properties of the hyperbolic tangent,

$$\Sigma_\tau = \sigma_f\gamma\coth\left[\coth^{-1}\left(\frac{\Sigma_0}{\sigma_f\gamma}\right) + \frac{\sigma_f}{\gamma}\tau\right] \quad \text{(A.23)}$$

Since coth is invertible in the positive domain and $\coth(+\infty) = 1$ 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:

$$\mathcal{V}_\tau \equiv \mathbb{E}_0[\hat{\mu}_\tau^2] = \mathbb{E}_0[\mu_\tau^2] + \mathbb{E}_0[(\hat{\mu}_\tau - \mu_t)^2] - 2\mathbb{E}_0[(\hat{\mu}_\tau - \mu_\tau)\mu_\tau] = \mathbb{E}_0[\mu_\tau^2] - \Sigma_\tau \quad \text{(A.24)}$$

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.22) and (A.23) into (A.24) and using $\Omega_\tau = \gamma\Sigma_\tau$, we get:

$$\mathcal{V}_\tau = \sigma_f^2\tau + \gamma\left(\Omega_0 - \Omega_\tau\right) = \sigma_f^2\tau + \gamma\left(\Omega_0 - \sigma_f \coth\left(arcoth\left(\frac{\Omega_0}{\sigma_f}\right) + \frac{\sigma_f}{\gamma}\tau\right)\right) = \sigma_f^2\tau + \mathcal{L}_\tau^{\mathcal{V}} \tag{A.25}$$

where we define the learning component as: $\mathcal{L}_\tau^{\mathcal{V}} \equiv \gamma\sigma_f\left(\frac{\Omega_0}{\sigma_f} - \coth\left(arcoth\left(\frac{\Omega_0}{\sigma_f}\right) + \frac{\sigma_f}{\gamma}\tau\right)\right)$. Since coth is decreasing and convex, $V_{\tau(\Omega_0)}$ is increasing and concave in τ and Ω_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 apply this 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:

$$h_\tau(\Omega_0) \equiv \frac{f(\tau|\Omega_0)}{\int_\tau^\infty f(s|\Omega)ds} = \frac{f(\mathcal{V}_\tau(\Omega_0), \bar{\mu}_0)}{\int_\tau^\infty f(\mathcal{V}_s(\Omega_0), \bar{\mu}_0)ds} = \frac{\mathcal{V}_\tau'(\Omega_0)\sum_{j=0}^\infty \alpha_j \exp\left(-\beta_j\frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)}{\int_\tau^\infty \mathcal{V}_s'(\Omega_0)\sum_{j=0}^\infty \alpha_j \exp\left(-\beta_j\frac{\mathcal{V}_s(\Omega_0)}{\bar{\mu}_0^2}\right)ds} \tag{A.26}$$

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.26):

$$h_\tau(\Omega_0) = \frac{\pi^2}{8\bar{\mu}_0^2}\Psi\left(\frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}\right)\mathcal{V}_\tau'(\Omega_0) \tag{A.27}$$

where we define $\Psi(x) \equiv \frac{\sum_{j=0}^\infty \alpha_j \exp\left(-\beta_j x\right)}{\sum_{j=0}^\infty \frac{1}{\alpha_j}\exp\left(-\beta_j x\right)}$ 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.

$\square$

Proof of Proposition 7 (Uncertainty Distributions). The strategy for the proof is as follows. We derive the Kolmogorov Forward Equation (KFE) of the joint ergodic distribution using the adjoint operator, and we find its zeros 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:

$$\mathcal{R}^\circ \equiv \{(\hat{\mu},\Omega) \in [-\infty,\infty] \times [\sigma_f,\infty] \quad \text{such that} \quad |\hat{\mu}| < \bar{\mu}(\Omega) \;\&\; \hat{\mu} \neq 0\} \tag{A.28}$$

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^* in (A.2) we write the KFE as:

$$-\frac{\sigma_f^2 - \Omega^2}{\gamma} f_\Omega(\hat{\mu},\Omega) + \frac{2\Omega}{\gamma} f(\hat{\mu},\Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu},\Omega) + \lambda\left[f\left(\hat{\mu},\Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu},\Omega)\right] = 0 \tag{A.29}$$

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:

$$\frac{2\Omega}{\gamma} f(\hat{\mu},\Omega) + \frac{\Omega^2 - \Omega^{*2}}{\gamma} f_\Omega(\hat{\mu},\Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu},\Omega) = 0 \tag{A.30}$$

with two border conditions:

$$\forall\Omega \quad f(|\bar{\mu}(\Omega)|,\Omega) = 0 \quad ; \quad \int_{\Omega\geq\sigma_f}\int_{|\hat{\mu}|\leq\bar{\mu}(\Omega)} f(\hat{\mu},\Omega)d\hat{\mu}d\Omega = 1 \tag{A.31}$$

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:

$$f(\hat{\mu},\Omega) = h(\Omega)g(\hat{\mu},\Omega) \tag{A.32}$$

Substituting (A.32) into (A.30) 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.32) into (A.31):

$$\forall\Omega \quad h(\Omega)g(|\bar{\mu}(\Omega)|,\Omega) = 0 \quad ; \quad \int_{\Omega\geq\sigma_f}\int_{|\hat{\mu}|\leq\bar{\mu}(\Omega)} h(\Omega)g(\hat{\mu},\Omega)d\hat{\mu}d\Omega = 1 \tag{A.33}$$

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):

$$r(\Omega) \propto 2 f(\bar{\mu}(\Omega), \Omega) \frac{\sigma_f^2 - \Omega^2}{\gamma} + \lambda \int_{-\bar{\mu}\left(\Omega - \sigma_u^2/\gamma\right)}^{\bar{\mu}\left(\Omega - \sigma_u^2/\gamma\right)} f\left(\mu, \Omega - \frac{\sigma_u^2}{\gamma}\right) I\left(\hat{\mu} > \bar{\mu}\left(\Omega\right)\right) d\mu d\Omega + 2 |f_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega)| \frac{\Omega^2}{2} \tag{A.34}$$

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:

$$\frac{r(\Omega)}{h(\Omega)} \propto |g_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega)| \Omega^2 \tag{A.35}$$

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

$$g_{\mu^2}(\hat{\mu}, \Omega^*) = 0; \qquad g(\bar{\mu}(\Omega^*), \Omega^*) = 0; \qquad \int_{-\bar{\mu}(\Omega^*)}^{\bar{\mu}(\Omega^*)} g(\hat{\mu}, \Omega^*) d\hat{\mu} = 1 \qquad g \in \mathbb{C} \tag{A.36}$$

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 :

$$g(\mu, \hat{\Omega}) = \begin{cases} \frac{1}{2\bar{\mu}(\Omega^*)} \left(1 + \frac{\hat{\mu}}{\bar{\mu}(\Omega^*)} \right) & \text{if } \hat{\mu} \in [-\bar{\mu}(\hat{\Omega}), 0] \\ \frac{1}{2\bar{\mu}(\Omega^*)} \left(1 - \frac{\hat{\mu}}{\bar{\mu}(\Omega^*)} \right) & \text{if } \hat{\mu} \in (0, \bar{\mu}(\hat{\Omega})] \end{cases} \tag{A.37}$$

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

6. **Ratio when $\Omega \approx \Omega^*$.** By the previous result and using continuity, the ratio of the renewal to marginal distributions near Ω^* is equal to:

$$\frac{r(\Omega)}{h(\Omega)} = |g_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega)| \, \Omega^2 = \frac{\Omega^2}{2\bar{\mu}(\Omega)^2} = \frac{1}{2\mathbb{E}[\tau|(0, \Omega)]} \tag{A.38}$$

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

□

Proof of Proposition 8 (Uncertainty Heterogeneity and Price Statistics). 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 $r(\Omega)$, i.e. 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):

$$\int_0^\infty \mathbb{E}\left[\hat{\mu}_\tau^2 \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] r(\tilde{\Omega}) d\tilde{\Omega} = \mathbb{E}[(\Delta p)^2] = \mathbb{V}[(\Delta p)] \tag{A.39}$$

- Notice that by the law of large numbers, if time is sufficiently long, the mean uncertainty is the same as the average uncertainty of one firm for every initial condition:

$$\mathbb{E}[\Omega^2] = \lim_{T \to \infty} \frac{\int_0^T \Omega_s^2 ds}{T} \tag{A.40}$$

  Therefore

$$\begin{aligned}
\mathbb{E}[\Omega^2] &= \lim_{T \to \infty} \frac{\int_0^T \Omega_s^2 ds}{T} = \lim_{n \to \infty} \frac{n^{-1} \sum_{i=0}^n \int_{\tau_{i-1}}^{\tau_i} \Omega_s^2 ds}{n^{-1} \sum_{i=0}^n \int_{\tau_{i-1}}^{\tau_i} ds} = \frac{\lim_{n \to \infty} n^{-1} \sum_{i=0}^n \int_0^{\tau_i - \tau_{i-1}} \Omega_{\tau_{i-1}+s}^2 ds}{\lim_{n \to \infty} n^{-1} \sum_{i=0}^n \int_0^{\tau_i - \tau_{i-1}} ds} \\
&= \frac{\mathbb{E}\left[\int_0^\tau \Omega_s^2 ds\right]}{\mathbb{E}\left[\int_0^\tau ds\right]} = \frac{\int_0^\infty \mathbb{E}[\int_0^\tau \Omega^2 ds | (\mu_0, \Omega_0) = (0, \tilde{\Omega})] r(\tilde{\Omega}) d\tilde{\Omega}}{\int_0^\infty \mathbb{E}[\tau | (\mu_0, \Omega_0) = (0, \tilde{\Omega})] r(\tilde{\Omega}) d\tilde{\Omega}}
\end{aligned} \tag{A.41}$$

  where the denominator is equal to the mean expected time $\mathbb{E}[\tau]$. Therefore,

$$\int_0^\infty \mathbb{E}\left[\int_0^\tau \Omega_s^2 ds \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] r(\tilde{\Omega}) d\tilde{\Omega} \quad = \quad \mathbb{E}[\tau]\mathbb{E}[\Omega^2] \tag{A.42}$$

- Putting together (A.39) and (A.42) we get the first result:

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

We conclude by showing the relationship with long-run uncertainty Ω^*. Using properties of the compound Poisson process, we have that in the cross-section, $\mathbb{E}[\sigma_u^2 dQ_t] = \sigma_u^2 \lambda dt$. Therefore, the first and second comments of the population's average uncertainty are related in the following way

$$d\mathbb{E}[\Omega_t] = \frac{\sigma_f^2 + \lambda \sigma_u^2 - \mathbb{E}[\Omega_t^2]}{\gamma^2} dt = \frac{(\Omega^*)^2 - \mathbb{E}[\Omega_t^2]}{\gamma^2} dt \tag{A.43}$$

Since the mean of the ergodic distribution satisfies $d\mathbb{E}[\Omega_t] = 0$, we have that $(\Omega^*)^2 = \mathbb{E}[\Omega_t^2]$. □

Proof of Propositions 9,10, and 11 (Macroeconomic Effects of Uncertainty Cycles). This is a general proof that shows all the results in Proposition 9, 10, and 11. Assume the economy is in steady state and it is hit with one-time unanticipated monetary shock of size $\delta > 0$. Simultaneously, idiosyncratic uncertainty increases by $\kappa\mathbb{E}[\Omega]$ for all firms. If firms only observe a fraction $\alpha \in [0,1]$ of the monetary shock, then:

1. The total output effect $\mathcal{M}(\delta,\alpha,\kappa) \equiv \int_0^\infty \tilde{Y}_s ds$ is the sum of inaction and forecast errors:

$$
\begin{aligned}
\mathcal{M}(\delta,\alpha,\kappa) &= \mathcal{I}(\delta,\alpha,\kappa) + (1-\alpha)\delta\mathcal{F}(\kappa) && \text{(A.44)}\\
\mathcal{I}(\delta,\alpha,\kappa) &= -\int_{\hat{\mu},\Omega} w(\hat{\mu}-\alpha\delta, \Omega+\kappa\mathbb{E}[\Omega], (1-\alpha)\delta) dF(\hat{\mu},\Omega) && \text{(A.45)}\\
w(\hat{\mu},\Omega,\varphi) &= \mathbb{E}\left[\int_0^\tau \hat{\mu}_t \; dt + w(0,\Omega_\tau, \varphi e^{-\int_0^t \frac{\Omega_s}{\gamma}ds}) \Big| (\hat{\mu}_0,\Omega_0) = (\hat{\mu},\Omega)\right] && \text{(A.46)}\\
\mathcal{F}(\kappa) &= \int_\Omega \mathbb{E}\left[\int_0^\infty e^{-\int_0^t \frac{\Omega_s}{\gamma}ds} dt \Big| \Omega_0 = \Omega + \kappa\mathbb{E}[\Omega]\right] dh(\Omega) && \text{(A.47)}
\end{aligned}
$$

subject to the following stochastic processes for (biased) estimates and uncertainty:

$$
d\hat{\mu}_t = -\Omega_t \frac{e^{-\int_0^t \Omega_s/\gamma ds}}{\gamma} dt + \Omega_t d\hat{Z}_t; \quad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t
$$

2. The following properties hold:

   - **Disclosed shock** ($\alpha = 1$) **and no uncertainty shock** ($\kappa = 0$) Pricing mistakes after the first price change are equal to zero $w(0,\Omega,0) = 0$ and the total output response is given by

$$
\begin{aligned}
\mathcal{M}(\delta,1,0) &= \mathcal{I}(\delta,1,0) = -\int_{\hat{\mu},\Omega} w(\hat{\mu}-\delta,\Omega) dF(\hat{\mu},\Omega) && \text{(A.48)}\\
w(\hat{\mu},\Omega,0) &= \mathbb{E}\left[\int_0^\tau \hat{\mu}_t \; dt \Big| (\hat{\mu}_0,\Omega_0) = (\hat{\mu},\Omega)\right] && \text{(A.49)}
\end{aligned}
$$

   - **Undisclosed shock** ($\alpha \in [0,1)$) **and no uncertainty shock** ($\kappa = 0$) Up to a first order, the total output effect is given by

$$
\mathcal{M}(\delta,\alpha,0) \geq \delta\left(\frac{\alpha\mathbb{E}[\tau]}{6} + (1-\alpha)\underline{\mathcal{F}}\right) \tag{A.50}
$$

     where the RSH is the first order effect of a monetary shock without heterogeneity ($\lambda = 0$) and $\underline{\mathcal{F}} \equiv \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}$ is a function of price statistics.

   - **Undisclosed shock** ($\alpha \in [0,1)$) **and aggregate uncertainty shock** ($\kappa > 0$) For all $\kappa > 0$, the forecast errors $\mathcal{F}(\kappa)$ are bounded below by

$$
\mathcal{F}(\kappa) \geq \xi\left(\frac{\mathbb{E}[\tau]\mathbb{E}[\Omega]^2(1+\kappa)^2}{\mathbb{V}[\Delta p]}\right)\underline{\mathcal{F}} \geq \xi((1+\kappa)^2)\;\underline{\mathcal{F}} \tag{A.51}
$$

     where the expression in the last inequality refers to the case without heterogeneity ($\lambda = 0$). The function ξ is decreasing, $\lim_{x\to 1}\xi(x) = 1$ and $\lim_{x\to\infty}\xi(x) = 0$

□

Proof. We first identify the initial conditions and the stochastic processes for an augmented firm state that includes forecast errors: $(\hat{\mu}_t, \Omega_t, \varphi_t)$. Then we characterize recursively the two components of the output effect, the inaction errors and the forecast errors. Finally, we show the properties of the output effect for particular sets of parameters.

1. **State's initial conditions**

   - A positive monetary shock of size $\delta > 0$ 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 $\hat{\mu}_t - \mu_t = \tilde{\varphi}_t \sim N(0,\gamma\Omega_t)$. 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: $\tilde{\varphi}_0 \sim \mathcal{N}((1-\alpha)\delta, \gamma\Omega_0)$. Notice that we can decompose $\tilde{\varphi}_0 = \varphi_0 + \hat{\varphi}_0^2$ with $\varphi_0 = (1-\alpha)\delta$ and $\hat{\varphi}_0 \sim \mathcal{N}(0,\gamma\Omega_0)$. We will continue with this decomposition for all t; a deterministic bias and the stochastic unbiased part of the forecast error.

- Finally, uncertainty gets amplified by a factor κ, thus $\Omega_0 = \Omega_{-1} + \kappa \mathbb{E}[\Omega]$.

2. **State's stochastic process** Now we derive the laws of motion for $(\hat{\mu}_t, \Omega_t, \varphi_t)$ for $t > 0$.

   - From equation (A.7), together with the definition of forecast errors, we can write the process for markup gap estimates in terms of forecast errors instead of the innovations' representation:

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

   where in the last step we used the same argument as in A.7.

   - By construction, the forecast error only has a drift term $d\varphi_t = X(\omega)dt$. By the definition of the forecast error and the stochastic process of markup gap, $0 = \mathbb{E}[d\mu_t] = \mathbb{E}[-d\hat{\varphi}_t + d\hat{\mu}_t] = -\mathbb{E}[d\hat{\varphi}_t] + \mathbb{E}[d\hat{\mu}_t] = -\mathbb{E}[d\hat{\varphi}_t] - \Omega_t \frac{\varphi_t}{\gamma} dt$ and thus we have that

   $$d\varphi_t = -\frac{\Omega_t}{\gamma} \varphi_t dt$$

   - The process for uncertainty is the same as (A.6).

3. **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:

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

   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:

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

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

4. **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

   $$\begin{aligned} \mathcal{M}(\delta, \alpha, \kappa) &\equiv -\int_0^{\infty} \tilde{Y}_t dt \\ &= -\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega} (\hat{\mu}_t + \tilde{\varphi}_t) \, dF_t(\hat{\mu}_t, \Omega_t, \tilde{\varphi}_t) \right] dt \\ &= -\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega} (\hat{\mu}_t + \varphi_t) \, dF_t(\hat{\mu}_t, \Omega_t, \varphi_t) \right] dt \\ &= \underbrace{-\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega} \hat{\mu}_t dF_t(\hat{\mu}_t, \Omega_t, \varphi_t) \right] dt}_{\text{average inaction error}} + \underbrace{\int_0^{\infty} \left[\int_{\hat{\mu}, \Omega} \varphi_t dF_t(\hat{\mu}_t, \Omega_t, \varphi_t) \right] dt}_{\text{average forecast error}} \end{aligned}$$

   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, and the forecast error can be decomposed between the mean and the unbiased part that cancels in the aggregate. The first term captures the average inaction error and the second term the average forecast error.

- Average inaction error. The average inaction error is computed as:

$$\begin{aligned}\mathcal{I}(\delta,\alpha,\kappa) &\equiv -\int_0^\infty \left[\int_{\hat{\mu},\Omega} \hat{\mu}_t dF_{t0}(\hat{\mu}_t,\Omega_t,\varphi_t|\hat{\mu}_0,\Omega_0,(1-\alpha)\delta)dF(\hat{\mu}_0,\Omega_0)\right] dt \\ &= -\int_{\hat{\mu},\Omega} w(\hat{\mu}_{-1}-\alpha\delta,\Omega_{-1}+\kappa\mathbb{E}[\Omega],(1-\alpha)\delta)dF(\hat{\mu}_{-1},\Omega_{-1})\end{aligned}$$

where we exchange the expectation and integral operators, and use the definition of w.

- Average forecast error. Since $d\varphi_t = -\frac{\Omega_t}{\gamma}dt$, we have that $\varphi_t = (1-\alpha)\delta e^{-\int_0^t \frac{\Omega_s}{\gamma}ds}$, thus the average forecast error is equal to $(1-\alpha)\delta\mathcal{F}(\kappa)$ where

$$\mathcal{F}(\kappa) \equiv \int_{\Omega_{-1}} \mathbb{E}\left[\int_0^\infty e^{-\int_0^t \frac{\Omega_s}{\gamma}ds}dt|\Omega_0 = \Omega_{-1}+\kappa\mathbb{E}[\Omega]\right] dh(\Omega)$$

5. **Disclosed shock** ($\alpha = 1$) **and no uncertainty shock** ($\kappa = 0$). By the symmetry of the Brownian motion, we have that $w(0,\kappa,0) = 0$. Therefore, for $\alpha = 0$, we have that we only need to keep track of the first price change:

$$w(\hat{\mu},\Omega,0) = \mathbb{E}\left[\int_0^\tau \hat{\mu}_t\ dt\Big|(\hat{\mu}_0,\Omega_0) = (\hat{\mu},\Omega)\right]$$

6. **Undisclosed shock** ($\alpha \in [0,1)$) **and no uncertainty shock** ($\kappa = 0$). We first derive a lower bound for the forecast errors and then for the average inaction error.

- Lower bound for forecast errors. Since $f(x) = e^x$ and $f(x) = x^2$ are convex functions, we have that

$$\begin{aligned}\int_{\Omega_{-1}} \mathbb{E}\left[\int_0^\infty e^{-\int_0^t \frac{\Omega_s}{\gamma}ds}dt|\Omega_0 = \Omega_{-1}\right] dh(\Omega) &\geq \int_0^\infty e^{-\int_0^t \frac{\int_{\Omega_{-1}}\mathbb{E}\left[\Omega_s|\Omega_0=\Omega_{-1}\right]dh(\Omega)}{\gamma}ds}dt \\ &= \int_0^\infty e^{-\frac{\mathbb{E}[\Omega]}{\gamma}t}dt = \frac{\gamma}{\mathbb{E}[\Omega]} \geq \frac{\gamma}{\sqrt{\mathbb{E}[\Omega^2]}} = \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}\end{aligned}$$

Therefore, we have that $\mathcal{F}(0) \geq \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}} = \underline{\mathcal{F}}$, where $\underline{\mathcal{F}} \equiv \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}$ is a function of price statistics. Our result in Proposition 8, applied to the case without heterogeneity ($\lambda = 0$) implies that $\gamma/\sigma_f = \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}$. Therefore, the lower bound in the RHS is equal to the average forecast error in the case without heterogeneity.

- Lower bound for inaction errors. We derive a lower bound for the average inaction errors in two steps. First, given that the policy is independent from the forecast error, and these only increase the inaction error, we consider the modified inaction errors $\mathcal{I}^*(\delta,\alpha,0)$ as follows:

$$\begin{aligned}\mathcal{I}^*(\delta,\alpha,0) &= -\int_{\hat{\mu},\Omega} w^*(\hat{\mu}-\alpha\delta,\Omega)dF(\hat{\mu},\Omega) \\ w^*(\hat{\mu},\Omega) &= \mathbb{E}\left[\int_0^\tau \hat{\mu}_t\ dt\Big|(\hat{\mu}_0,\Omega_0) = (\hat{\mu},\Omega)\right] \\ d\hat{\mu}_t &= \Omega_t d\hat{Z}_t; \quad d\Omega_t = \frac{\sigma_f^2-\Omega_t^2}{\gamma}dt + \frac{\sigma_u^2}{\gamma}dQ_t\end{aligned}$$

where $\mathcal{I}(\delta,\alpha,0) \geq \mathcal{I}^*(\delta,\alpha,0)$. The second step consists in showing that $\mathcal{I}^*(\delta,\alpha,0) \geq \delta\frac{\mathbb{E}[\tau]}{6}$. Notice that what matters for computation of output effects is the effective size of the shock, given by $\hat{\delta} \equiv \alpha\delta$. Therefore, $\mathcal{I}^*(\delta,\alpha,0) = \mathcal{I}^*(\hat{\delta},1,0)$. Recall that the joint distribution can be decomposed as $f(\hat{\mu},\Omega) = g(\hat{\mu}|\Omega)h(\Omega)$. Then we can rewrite

$$\mathcal{I}^*(1,\hat{\delta},0) = \int_\Omega \mathbb{E}[\tau|\Omega]v(\Omega,\hat{\delta})dh(\Omega), \quad where \quad v(\Omega,\hat{\delta}) = -\frac{\int_{\hat{\mu},\Omega} w^*(\hat{\mu}-\hat{\delta},\Omega)g(\hat{\mu}|\Omega)d\hat{\mu}}{\mathbb{E}[\tau|\Omega]}$$

$$= \mathbb{E}^h\left[\mathbb{E}[\tau|\Omega]\right]\mathbb{E}^h\left[\mathbb{E}[v(\Omega,\hat{\delta})|\Omega]\right] + \mathbb{C}ov^h\left[\mathbb{E}[\tau|\Omega], v(\Omega,\hat{\delta})\right] \tag{A.52}$$

$$\geq \mathbb{E}^h\left[\mathbb{E}[\tau|\Omega]\right]\mathbb{E}^h\left[\mathbb{E}[v(\Omega,\hat{\delta})|\Omega]\right] \tag{A.53}$$

$$\geq \mathbb{E}^r\left[\mathbb{E}[\tau|\Omega]\right]\mathbb{E}^h\left[\mathbb{E}[v(\Omega,\hat{\delta})|\Omega]\right] \tag{A.54}$$

$$\geq \left[\mathbb{E}[\tau]\right]\frac{\hat{\delta}}{6} \tag{A.55}$$

The function $v(\Omega,\hat{\delta})$ is the effect of the monetary shock conditional on uncertainty Ω. It can be verified that this function is decreasing in Ω. As $\Omega \to \infty$, price changes are mostly driven by the Brownian process affecting markup estimates (the drift generated by forecast errors and the Poisson shocks in uncertainty lose importance). Therefore, we apply the result in Álvarez, Le Bihan and Lippi (2014) for a shock of size $\hat{\delta}$ and obtain: $v(\Omega,\hat{\delta}) \geq \lim_{\Omega\to\infty} v(\Omega,\hat{\delta}) = \frac{\hat{\delta}}{6}$. Summarizing, the inaction errors are bounded below by: $\mathcal{I}(\delta,\alpha,0) \geq \frac{\alpha\delta\mathbb{E}[\tau]}{6}$. With these results we can justify each step in (A.55): the second equality is due to the definition of covariance, the third inequality is due to the positive relation between $\mathbb{E}[\tau|\Omega], v(\Omega,\hat{\delta})$, the third inequality is since the renewal distribution puts more weight to high uncertainty and the last step is since $v(\Omega,\hat{\delta}) \geq \lim_{\Omega\to\infty} v(\Omega,\hat{\delta}) = \frac{\hat{\delta}}{6}$.

7. **Undisclosed shock** ($\alpha \in [0,1)$) **and aggregate uncertainty shock** ($\kappa > 0$). Now we characterize an upper bound of the rate of converge of the forecast error after a monetary shock.

   - First notice that

$$\mathcal{F}(\kappa) \geq \int_0^\infty e^{-\int_0^t \frac{\int_{\Omega_{-1}} \mathbb{E}\left[\Omega_s|\Omega_0=\Omega_{-1}+\kappa\mathbb{E}[\Omega]=\right]dh(\Omega)}{\gamma}ds}dt = \int_0^\infty e^{-\int_0^t \frac{y_s}{\gamma}ds}dt$$

   where we define $y_s \equiv \int_{\Omega_{-1}} \mathbb{E}\left[\Omega_s|\Omega_0=\Omega_{-1}+\kappa\right]dh(\Omega) = \mathbb{E}_i[\Omega_s^i]$. Using uncertainty's law of motion together with the result in Proposition 8, we have that

$$dy_s = \mathbb{E}_i\left[d\Omega_s^i\right] = \mathbb{E}\left[\frac{\sigma_f^2-{\Omega_s^i}^2}{\gamma}ds + \frac{\sigma_u^2}{\gamma}dQ_s^i\right] = \gamma^{-1}\left(\sigma_f^2+\lambda\sigma_u^2-\mathbb{E}\left[{\Omega_s^i}^2\right]\right)ds = \gamma^{-1}\left(\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}-\mathbb{E}_i\left[{\Omega_s^i}^2\right]\right)ds$$

   with initial condition $y_0 = \mathbb{E}[\Omega](1+\kappa)$. Since $\mathbb{E}_i[{\Omega_s^i}^2] \geq \mathbb{E}_i[\Omega_s^i]^2$ by Jensen's inequality, we obtain

$$dy_s = \gamma^{-1}\left(\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}-\mathbb{E}_i[{\Omega_s^i}^2]\right)ds \leq \gamma^{-1}\left(\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}-y_s^2\right)ds$$

   - Let $\tilde{y}_t$ be the solution to $d\tilde{y}_s = \gamma^{-1}\left(\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}-\tilde{y}_s^2\right)ds$ with initial condition $\tilde{y}_0 = y_0$. Since $dy_s \leq d\tilde{y}_t$ for all t and they start at the same value, $\tilde{y}_s$ is always above y_s and we get

$$\mathcal{F}(\kappa) \geq \int_0^\infty e^{-\int_0^t \frac{\tilde{y}_s}{\gamma}ds}dt$$

   where the exact solution to $\tilde{y}_s$ is given by

$$\tilde{y}_s = \sqrt{\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}}\coth\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)+\sqrt{\frac{\mathbb{V}[\Delta p]}{\gamma^2\mathbb{E}[\tau]}}s\right)$$

   - Therefore, forecast errors are bounded below by:

$$\begin{aligned}
\mathcal{F}(\kappa) &\geq \int_0^\infty \exp\left(-\int_0^t \frac{\tilde{y}_s}{\gamma}ds\right)dt \\
&= \int_0^\infty \exp\left(-\int_0^t \frac{\sqrt{\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}}\coth\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)+\sqrt{\frac{\mathbb{V}[\Delta p]}{\gamma^2\mathbb{E}[\tau]}}s\right)}{\gamma}ds\right)dt \\
&= \int_0^\infty \exp\left(-\int_0^{\sqrt{\frac{\mathbb{V}[\Delta p]}{\gamma^2\mathbb{E}[\tau]}}t}\coth\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)+u\right)du\right)dt \\
&= \int_0^\infty \exp\left(-\log\left(\sinh\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)+t\right)\right)\Bigg|_0^{\sqrt{\frac{\mathbb{V}[\Delta p]}{\gamma^2\mathbb{E}[\tau]}}t}\right)dt
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \frac{\sinh\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)+\sqrt{\frac{\mathbb{V}[\Delta p]}{\gamma^2\mathbb{E}[\tau]}}t\right)}{\sinh\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)\right)}dt \\
&= \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}\int_0^\infty \frac{\sinh\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)+s\right)}{\sinh\left(\coth^{-1}\left(\sqrt{\frac{\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}y_0\right)\right)}ds \\
&= \xi\left(\frac{\mathbb{E}[\tau]\mathbb{E}[\Omega]^2(1+\kappa)^2}{\mathbb{V}[\Delta p]}\right)\underline{\mathcal{F}}
\end{aligned}
$$

where $\underline{\mathcal{F}} \equiv \sqrt{\frac{\gamma^2\mathbb{E}[\tau]}{\mathbb{V}[\Delta p]}}$ and the function ξ are defined also in terms of steady state price statistics. The function ξ is decreasing, $\lim_{x\to 1}\xi(x) = 1$ and $\lim_{x\to\infty}\xi(x) = 0$. Summarizing, $\mathcal{F}(\kappa) \geq \xi\left(\frac{\mathbb{E}[\tau]\mathbb{E}[\Omega]^2(1+\kappa)^2}{\mathbb{V}[\Delta p]}\right)\underline{\mathcal{F}}$.

- Finally, when $\lambda = 0$, the lower bound collapses to $\xi\left((1+\kappa)^2\right)\underline{\mathcal{F}}$.

$\square$

Proof of Proposition 12 (Revenue losses due to frictions). Let $\Pi(\upmu_t)$ be the flow profits as a function of the markups $\upmu_t$. A second order approximation to the profit function around the optimal frictionless markup $\upmu^*$ yields

$$
\Pi(\upmu_t) = \Pi(\upmu^*) + \frac{\Pi''(\upmu^*)}{2}(\upmu^*)^2\left(\frac{\upmu_t-\upmu^*}{\upmu^*}\right)^2 = \Pi(\upmu^*) + \frac{\Pi''(\upmu^*)}{2}(\upmu^*)^2\mu_t^2, \tag{A.56}
$$

where $\mu_t \equiv \log(\upmu_t/\upmu^*)$ is the realized markup-gap. Given the CES demand and the constant returns to scale technology, we can express the expected losses that arise from frictions (both nominal and informational) relative to the frictionless case, expressed as a fraction of revenue, are given by:

$$
\Delta \equiv \mathbb{E}\left[\frac{\Pi(\upmu_t)-\Pi(\upmu^*)}{R(\upmu^*)}\right] = \frac{1}{2}\left(\frac{\Pi''(\upmu^*)\,(\upmu^*)^2}{\Pi(\upmu^*)}\right)\frac{\Pi(\upmu^*)}{R(\upmu^*)}\mathbb{E}\left[\mu_t^2\right] = \frac{1}{2}\eta(1-\eta)\frac{1}{\eta}\mu_t^2 = \frac{1-\eta}{2}\mathbb{E}\left[\mu_t^2\right] \tag{A.57}
$$

To characterize the expectation, note that for each firm z:

$$
\begin{aligned}
\mathbb{E}\left[\mu_t(z)^2\right] &= \mathbb{E}\left[(\mu_t(z)-\hat{\mu}_t(z)+\hat{\mu}_t(z))^2\right] &\text{(A.58)}\\
&= \mathbb{E}\left[(\mu_t(z)-\hat{\mu}_t(z))^2 + 2\left(\mu_t(z)-\hat{\mu}_t(z)\right)\hat{\mu}_t(z) + \hat{\mu}_t^2(z)\right] &\text{(A.59)}\\
&= \underbrace{\mathbb{E}\left[(\mu_t(z)-\hat{\mu}_t(z))^2\right]}_{=\gamma\mathbb{E}[\Omega_t(z)]} + \underbrace{2\mathbb{E}\left[\left(\mu_t(z)-\hat{\mu}_t(z)\right)\hat{\mu}_t(z)\right]}_{=0} + \underbrace{\mathbb{E}\left[\hat{\mu}_t(z)^2\right]}_{=\mathbb{V}[\hat{\mu}_t(z)]} &\text{(A.60)}\\
&= \gamma\mathbb{E}\left[\Omega_t(z)\right] + \mathbb{V}\left[\hat{\mu}_t(z)\right], &\text{(A.61)}
\end{aligned}
$$

For the first term, we have substituted the definition of uncertainty, for the second term, we have used the law of the iterated expectations to show that the average forecast error in the population is equal to zero (since each estimate is unbiased), and for the third term, we have used that the average markup estimate is equal to zero. Therefore the total expected profit losses are given by

$$
\Delta = \frac{1-\eta}{2}\left[\underbrace{\gamma\mathbb{E}\left[\Omega_t\right]}_{\text{Loss from info. friction}} + \underbrace{\mathbb{V}\left[\hat{\mu}_t\right]}_{\text{Loss from pricing friction}}\right]. \tag{A.62}
$$

Finally, to obtain an upper bound for the term related to information frictions, note that using the Jensen inequality and Proposition 8 we have that

$$
\gamma\mathbb{E}\left[\Omega_t\right] \leq \gamma\sqrt{\mathbb{E}\left[\Omega_t^2\right]} = \gamma\sqrt{\frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]}} \tag{A.63}
$$

$\square$