Menu Costs, Uncertainty Cycles, and the Propagation of Nominal Shocks*

Isaac Baley† **Julio A. Blanco** ‡

CLICK HERE FOR LATEST VERSION

April 30, 2015

Abstract

Nominal shocks have long lasting effects on real economic activity, beyond those implied by the average frequency of price adjustment in micro data. This paper develops a price-setting model that explains this gap through the interplay of menu costs and uncertainty about productivity. Uncertainty arises from firms' inability to distinguish between permanent and transitory changes in their idiosyncratic productivity. Upon the arrival of a productivity shock, a firm's uncertainty spikes up and then fades with learning until the arrival of the next shock. These uncertainty cycles, when paired with menu costs, generate recurrent episodes of high frequency of price adjustment followed by episodes of low frequency of adjustment at the firm level. This time variation in the individual adjustment frequency is consistent with empirical patterns, in particular a decreasing hazard rate of adjustment, and it is key to understand the sluggish propagation of nominal shocks.

JEL: D8, E3, E5

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

*Previously circulated as "Learning to Price". We are especially thankful to Virgiliu Midrigan and Laura Veldkamp for their advice and to three anonymous referees for their constructive comments. We also thank Fernando Álvarez, Rudi Bachmann, Anmol Bhandari, Jarda Borovička, Katka Borovičková, Olivier Coibion, Mark Gertler, Ricardo Lagos, John Leahy, Francesco Lippi, Robert E. Lucas, Rody Manuelli, Simon Mongey, Joseph Mullins, Emi Nakamura, Gastón Navarro, Ricardo Reis, Thomas Sargent, Edouard Schaal, Ennio Stacchetti, Venky Venkateswaran, as well as seminar participants at 4th Ifo Conference on Macroeconomics and Survey Data 2013, Midwest Economics Association 2013, Society of Economic Dynamics 2013, New York University, Princeton University, Washington University St. Louis, St. Louis Fed, ASSA Meetings 2015, Federal Reserve Board, University of Toronto, Einaudi Institute, CREI, Pompeu Fabra, Bank of International Settlements, Singapore Management University, Carnegie Mellon, UC Davis, University of Melbourne, University of Sydney, Banco de México, and ITAM for very useful comments and suggestions. Julio A. Blanco gratefully acknowledges the hospitality of the St. Louis Fed where part of this paper was completed.

†New York University. isaac.baley@nyu.edu; https://files.nyu.edu/ib572/public

‡New York University. jab772@nyu.edu; https://files.nyu.edu/jab772/public

1 Introduction

How do nominal rigidities and information frictions in firms' pricing decisions influence the propagation of nominal shocks? This paper addresses this classic question in monetary economics by developing a price-setting model that combines the most salient features of two vast literatures on price-setting: menu cost models and imperfect information models. The interplay between these frictions can simultaneously explain micro price-settings facts and macro evidence on the sluggish propagation of nominal shocks.

Each friction alone cannot account for all the facts. On one hand, the advantage of menu costs models is that the frequency of adjustment is endogenous which makes them successful in matching micro evidence on price changes; however, those models produce modest non-neutrality unless they are coupled with other exogenous sources of persistence such as strategic complementarities, as in **?**. On the other hand, the advantage of models with information frictions is that they produce large nominal non-neutralities by imposing constraints on the information sets, as in **?**, **?** and **?**; however, firms adjust their prices every period which contradicts the frequency of adjustment in micro data.

The key contribution of this paper is to show that incorporating infrequent shocks of unknown magnitude in a menu cost model can simultaneously explain micro and macro price-setting facts. The starting point is the framework in **?** where firms receive noisy signals about their productivity and face a menu cost to adjust their prices[1]. As firms estimate their optimal prices, their decision rules become dependent on estimation variance or *uncertainty*. The new element is the "noisy" arrival of large infrequent shocks: firms know when a shock occurrs, but not its sign or size. Under this assumption, the infrequent *first* moment shocks paired with the noisy signals give rise to *second* moment shocks or uncertainty shocks. These shocks in turn affect the decision to adjust prices and the response to nominal shocks.

Large, infrequent idiosyncratic shocks were first introduced in menu cost models by **?** and then used by **?** as a way to account for the empirical patterns of pricing behavior. In our model, when an infrequent shock arrives, regardless of it is positive or negative, it generates a spike in uncertainty. This noisy arrival of infrequent shocks is interpreted as changes in the economic environment that affect firms' optimal prices but that they cannot assign a sign or magnitude to. Examples include product turnover, changes in the supply chain, innovation, changes in the fiscal and regulatory environment, new technologies and access to new markets. These shocks have the potential to push the optimal price of the firm either upwards or downwards; in expectation, firms think these changes have no effect on their optimal price, but their estimates do become more uncertain.

[1] In **?** firms pay an observation cost to get the truth revealed; the special case considered here makes this cost infinite and thus the true state is never revealed.

The main consequence of introducing infrequent unknown shocks is that the information friction remains active in steady state. Upon the arrival of an infrequent shock, firm's uncertainty about future productivity jumps up on impact and subsequently decreases gradually as a result of learning. Uncertainty will jump up again with the arrival of a new productivity shock. This mechanism implies a cross-sectional distribution of uncertainty that translates into heterogeneous decision rules. In particular, firms with high levels of uncertainty have a higher frequency of adjustment than firms with low levels of uncertainty. Without the infrequent shocks, uncertainty becomes constant and the heterogeneity in price-setting is eliminated, as in the model by **?**.

Since the information friction remains active in steady state, individual prices respond sluggishly to unanticipated nominal shocks; this in turn slows down the response of the aggregate price level. These aggregate implications are studied in a general equilibrium model where a continuum of firms solve the price-setting problem with menu costs and information frictions. The model is then calibrated using micro price statistics; in particular, the slope of the hazard rate of price changes is used to calibrate the signal noise. The results are as follows: the half life of output after a nominal shock is 50% larger than in a menu cost model without information frictions; this figure goes up to 300% if the infrequent shocks are also absent.

The uncertainty shocks in this paper contrast with the stochastic volatility process for productivity used elsewhere. **?** models stochastic volatility as an unsynchronized process across firms; their calibration with large menu costs produces near-neutrality even for the average level of volatility. **?** models countercyclical stochastic volatility, synchronized across all firms, to explain why monetary neutrality increases in recessions. In our paper the volatility of productivity shocks is fixed, and the idiosyncratic uncertainty shocks arise as the result of firms' learning mechanism. Even though shocks are unsynchronized, non-neutrality is high for average levels of uncertainty, in contrast with the first paper. Furthermore, as we explain below, non-neutrality is time varying when considering a correlated shock across firms, as in the second paper.

To dissect the mechanism at play, the output response is decomposed into two components. The first is the well-known selection effect, which is a measure of nominal rigidities; the second component is a learning effects, which consists of average forecast errors across firms, a measure of information frictions. This decomposition makes evident the ability of information frictions to increase the effect of monetary policy, and more importantly, that the presence of forecast errors is a testable implication of the theory. Nominal shocks make their way into estimated optimal prices very gradually and eventually will be completely incorporated into prices. As a result, forecast errors rise on impact and then converge back to zero. **?** provide empirical support for this type of delayed response of forecast errors to aggregate shocks.

There are two salient predictions of the uncertainty cycles in terms of the inaction region and price statistics. A first prediction is that higher uncertainty increases the the frequency of adjustments. As in any problem with fixed adjustment costs, the decision rule takes the form of an inaction region, but with information frictions the inaction region becomes a function of uncertainty. An increase in uncertainty generates two well-known opposing effects: the option value of waiting or "wait-and-see" effect becomes larger and makes the inaction region wider; at the same time, there is "volatility" effect that increases the probability of leaving the inaction region because firms are hit on average with larger shocks. The model produces a new "learning" effect that works in the same direction of the volatility effect and thus amplifies it. When uncertainty is high, firms increase the weight they put on new information when making their estimations. Since new information comes with noise, estimates become even more volatile and more price changes are triggered. For small adjustment costs, the volatility and learning effects dominate the option value effect; thus higher uncertainty is associated with a higher frequency of adjustment. There are at least two pieces of evidence of this mechanism. **?** documents a strong comovement of the cross-sectional dispersion of the price changes and the frequency of price adjustment in BLS data. **?** documents a positive relationship between the variance of firm-specific expectation errors and the frequency of adjustment using micro data from German firms.

A second prediction of the model is a decreasing hazard rate of adjustment. As a result of learning, firm uncertainty falls with time and the probability of adjusting also decreases. Consequently, a firm that just received an infrequent shock expects to transition from high uncertainty and frequent adjustment to low uncertainty and infrequent adjustments. The mechanism predicts a decreasing hazard rate of price adjustment based entirely on a learning mechanism[2]. Decreasing hazard rates have been documented in various datasets[3], and we use this empirical counterpart to discipline the information structure. The slope of the hazard rate is driven by the noise volatility, and therefore the empirical hazard rate can be used to assess the importance of the information friction. This approach of using a price statistic to recover information parameters was first suggested in **?**, and **?** uses it to calibrate a signal-noise ratio in a labor market framework.

Besides the large and persistent effects of nominal shocks, the model also predicts time-variation in the propagation of nominal shocks: in highly uncertain times prices react faster to these shocks and there are lower output effects. The quantitative exercise suggests that a nominal shock paired with a correlated uncertainty shock across firms sharply reduces the output effect: an uncertainty shock of one times its steady state level cuts output response by half, while an uncertainty shock of four times its steady state level (comparable to the magnitudes suggested in **?** and **?**) completely mutes the output response. Recent empirical evidence of this relationship can be found in **?**

[2]Alternative explanations for decreasing hazard rates are heterogeneity and survivor bias as noted in **?**, sales as in **?**, mean reverting shocks as in **?**, or experimentation as in **?**.

[3]Decreasing hazard rates are documented by **?** using monthly BLS data for consumer and producer prices, controlling for heterogeneity and sales, **?** using weekly scanner data, **?** using Dominik's weekly scanner data, and **?** using monthly CPI data for Euro zone countries.

studies the propagation of monetary shocks and its interaction with commonly used measures of uncertainty. Another study is **?** which documents a positive relationship between exchange-rate pass-through and dispersion of price changes, where dispersion may be generated by uncertainty at firm level.

The theory developed in this paper may be prove useful beyond price-setting frameworks. Models with non-convex adjustment costs, such as portfolio choice, investment, and firm entry and exit, can be extended to incorporate information frictions using the tools developed here.

2 Firm problem with nominal rigidities and information frictions

The model developed here combines an inaction problem arising from a non-convex adjustment cost together with a signal extraction problem, where the underlying state process features both continuous and infrequent shocks. The decision rule takes the form of an inaction region that varies and reflects the uncertainty dynamics. Furthermore, uncertainty never stabilizes due to the recurrent nature of the large infrequent shocks. Although the focus is on pricing decisions, the model is easy to generalize to other settings.

2.1 Environment

Consider a profit maximizing firm that chooses the price p_t at which to sell her product. She must pay a menu cost θ in units of product every time she changes the price. Let p_t^* be the price that the firm would set in the absence of any frictions; such price makes her markup (price over marginal cost) constant. This target price is not observed perfectly; only noisy signals are available to the firm. The firm has to choose the timing of the adjustments as well as the new reset prices. Time is continuous and the firm discounts the future at a rate r.

Quadratic loss function Let μ_t be the markup gap, or the log difference between the current markup and the unconstrained (constant) markup. Firms incur an instantaneous quadratic loss as the markup gap moves away from zero:

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

where $B > 0$ is a constant. Quadratic profit functions are standard in price setting models, for example in **?** and **?**, and can be motivated as second order approximations of more general profit functions as in **?**.

Stochastic process for markups The markup gap μ_t follows a jump-diffusion process

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

where W_t is a Wiener process, 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, where $u_t \sim \mathcal{N}(0,1)$. The process q_t is independent of W_t and u_t. Analogously, the markup gap can be also expressed as

$$\mu_t = \sigma_f W_t + \sigma_u \sum_{\kappa=0}^{q_t} u_\kappa$$

where $\{\kappa\}$ are the dates when $dq_\kappa = 1$ and $\sum_{\kappa=0}^{q_t} u_\kappa$ is a compound Poisson process. Note that $\mathbb{E}[\mu_t] = 0$ and $\mathbb{V}[\mu_t] = (\sigma_f^2 + \lambda\sigma_u^2)t$. It is assumed that $\sigma_f < \sigma_u$. **?** studies a version of this jump-diffusion process in the context of stock option valuations.

This process for markup gaps nests two specifications that are benchmarks in the literature:

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

ii) large *infrequent* shocks modeled through the Poisson process q_t with large volatility σ_u. These shocks produce a leptokurtic distribution of price changes and are used in **?** and **?** as a way to capture relevant features of the empirical price change distribution, such as fat tails.

Signals Firms do not observe their markup gaps directly. They get continuous noisy observations of them denoted by s_t. These noisy signals of the markup gap evolve according to

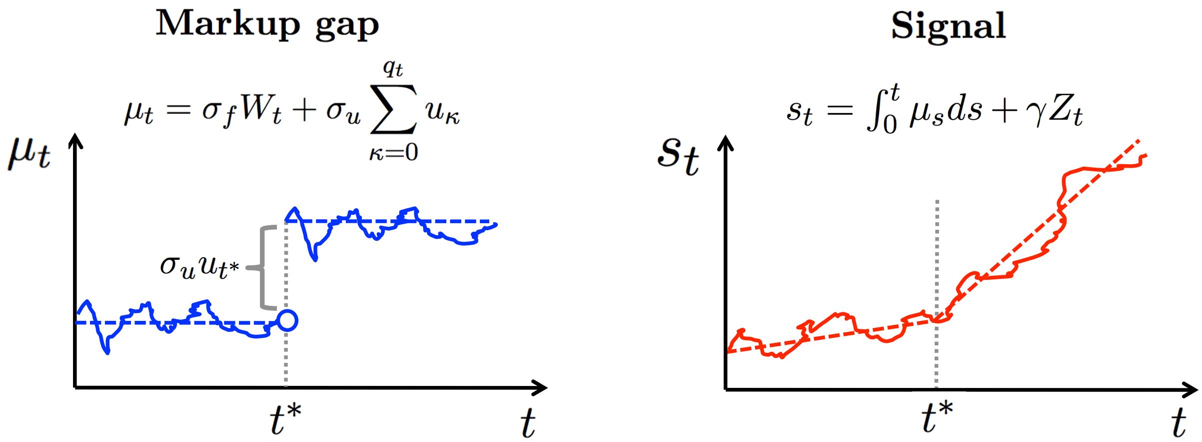
$$ds_t = \mu_t dt + \gamma dZ_t \tag{2}$$

where the signal noise Z_t follows a Wiener process, independent from W_t. The volatility parameter γ measures the size information friction. Note that the underlying state, μ_t, enters as the drift of the signal. This representation makes the filtering problem tractable (see **?** for details about filtering problems in continuous time). The signal process has continuous paths. To see this, rewrite it as $s_t = \int_0^t \mu_s ds + \gamma Z_t$ which is the sum of an integral and a Wiener process, and therefore it is continuous.

[4] **?** use a mean reverting process for productivity instead of a random walk. Still, our results concerning small frequent shocks will be compared with their setup.

Figure I illustrates the evolution of the markup gap and the signal process. When an infrequent shock arrives ($dq_t = 1$), the average *level* of the markup gap jumps to a new value; nevertheless, the signal has continuous paths and only its *slope* changes to a new average value.

Figure I:
llustration of the process for the markup gap and the signal.



Dashed lines illustrate a situation in which the Brownian motions F_t, T_t are not active and only infrequent shocks are present. The arrival of an infrequent shock changes the level of the markup gap and the slope of the signal.

Information set The information set at time t includes the history of signals $\tilde{\mu}$ as well as the realizations of the Poisson counter q, i.e. the firm knows if there has been an infrequent shock, but not the size of the innovation u_t. Thus the information set is given by the σ-algebra generated by the histories of these processes:

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

The assumption that the firm knows the arrival of infrequent shocks is made for analytical tractability. A more general approach would be to include an additional state variable to capture the hidden state q_t, although in that case the model must be solved numerically. This latter approach is used in hidden state Markov models pioneered by **?**.

2.2 Filtering problem

This section describes the filtering problem and derives the laws of motion for estimates and estimation variance. Equations (1) and (2) above jointly describe the noisy signals model for markup gaps, and are repeated here for convenience:

$$\begin{aligned} (state) \quad d\mu_t &= \sigma_f dW_t + \sigma_u u_t dq_t \\ (signal) \quad ds_t &= \mu_t dt + \gamma dZ_t \\ where \quad W_t, Z_t \sim \text{Brownian Motion}, \quad q_t &\sim Poisson(\lambda), \quad u_t \sim \mathcal{N}(0,1) \end{aligned}$$

Let $\hat{\mu}_t \equiv \mathbb{E}[\mu_t | I_t]$ be the best estimate of the markup gap (in a mean-squared error sense) and $\Sigma_t \equiv \mathbb{E}[(\mu_t - \hat{\mu}_t)^2 | I_t]$ 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. Finally, the innovation process X_t that represents the innovations or unpredictable component defined as the difference between the realization and its expectation plus the measurement noise:

$$(innovation) \qquad dX_t = \frac{1}{\gamma}(ds_t - \hat{\mu}_t dt) = \frac{1}{\gamma}(\mu_t - \hat{\mu}_t)dt + dZ_t \tag{3}$$

The innovation process is a one-dimensional Brownian motion under the probability distribution of the firm[5]. Proposition 1 below establishes the laws of motion for estimates and uncertainty. The derivation of these laws of motion involve an application of the Kalman-Bucy filter generalized to include the infrequent shocks.

Proposition 1 (Filtering equations) *Given the processes in (1), (2), and (3) and initial conditions $(\hat{\mu}_0, \Omega_0)$, the estimate of the markup gap $\hat{\mu}_t \equiv \mathbb{E}[\mu_t | I_t]$ and its uncertainty $\Omega_t \equiv \frac{\mathbb{E}[(\mu_t - \hat{\mu}_t)^2 | I_t]}{\gamma}$ evolve according to:*

$$d\hat{\mu}_t = \Omega_t dX_t \tag{4}$$

$$d\Omega_t = \underbrace{\frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt}_{deterministic} + \underbrace{\frac{\sigma_u^2}{\gamma} dq_t}_{uncertainty\ shocks} \tag{5}$$

where X_t is a Brownian motion.

Equation (4) says that the estimates $\hat{\mu}_t$ follow a Brownian motion X_t, given by the innovation process, with time varying volatility Ω_t. A notable characteristic of this filtering problem is that point estimates, as well as the signals and innovations, have continuous paths even though the underlying state is discontinuous. The continuity of these paths comes from the fact that changes in the state affect the slope of the innovations and signals but not their levels, and that the

[5] See Lemma 6.2.6 in **?** for a proof.

expected size of an infrequent shock u_t is zero. Markup estimations are not affected by the arrival of infrequent shocks, only uncertainty features jumps.

Higher uncertainty, more volatile estimates and faster learning A consequence of Bayesian learning is that when prior uncertainty is high, the estimates put more weight on the signals than on the previous estimate. To see how this is embedded in Equation (4) let's take a small period of time Δ. As a discrete process, the markup gap estimate at time t is given by the convex combination of the previous estimate $\hat{\mu}_{t-\Delta}$ and the signal $\tilde{\mu}_{t-\Delta}$, where the latter is the sum of the state and a white noise ϵ_t, and the weights are a function of prior uncertainty $\Omega_{t-\Delta}$ and signal noise γ:

$$\hat{\mu}_t = \underbrace{\frac{\gamma}{\Omega_{t-\Delta}\Delta + \gamma}}_{\text{weight on prior estimate}} \hat{\mu}_{t-\Delta} + \underbrace{\left(1 - \frac{\gamma}{\Omega_{t-\Delta}\Delta + \gamma}\right)}_{\text{weight on signal}} \left(\mu_t \Delta + \gamma \sqrt{\Delta} \epsilon_t\right), \quad \epsilon_t \sim \mathcal{N}(0,1) \tag{6}$$

For high prior uncertainty $\Omega_{t-\Delta}$, the weight that the forecast assigns to the signal is also high. This means that the estimate incorporates more information about the current markup μ_t; in other words learning is faster, but it also brings more white noise ϵ_t into the estimation.

Uncertainty cycles Regarding the evolution of uncertainty, Equation (5) shows that it is composed of a deterministic and a stochastic component, where the latter is active whenever the markup gap receives an infrequent shock. Let's study each component separately. In the absence of infrequent shocks ($\lambda = 0$), uncertainty Ω_t follows a deterministic path which converges to the constant volatility of the continuous shocks σ_f, i.e. the fundamental volatility of the markup gap. This case is studied in the Online Appendix of **?**. The deterministic convergence is a result of the learning process: as time goes by, estimation precision increases until the only volatility left is fundamental.

In the model with infrequent shocks ($\lambda > 0$), uncertainty jumps up on impact with the arrival of every infrequent shock and then decreases deterministically until the arrival of a new infrequent shock that will push uncertainty up again. The time series profile of uncertainty features a saw-toothed profile that never stabilizes due to the recurrent nature of these shocks. If the arrival of the infrequent shocks were not known and instead the firm had to filter their arrival as well, uncertainty would probably feature a hump-shaped profile instead of a jump.

Although uncertainty never settles down, it is convenient to characterize the level of uncertainty such that its expected change is equal to zero. This level of uncertainty is equal to the variance of the state $\mathbb{V}[\mu_t] = \Omega^{*2} t$, hence it is called "fundamental" uncertainty. The following Lemma establishes its value Ω^*.

Lemma 1 (Fundamental uncertainty) *If* $\Omega_t = \Omega^* \equiv \sqrt{\sigma_f^2 + \lambda \sigma_u^2}$, *then* $\mathbb{E}\left[\frac{d\Omega_t}{dt}\middle|\mathcal{I}_t\right] = 0$.

The next sections show that the ratio of current to fundamental uncertainty is a key determinant of decision rules and price statistics.

2.3 Decision rules

With the characterization of the filtering problem at hand, this section proceeds to characterize the price adjustment decision of the firm.

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

$$\max_{\{\mu_{\tau_i},\tau_i\}_{i=1}^{\infty}} -\mathbb{E}\left[\sum_{i=0}^{\infty} e^{-r\tau_{i+1}}\left(\theta + \int_{\tau_i}^{\tau_{i+1}} e^{-r(s-\tau_{i+1})} B\mu_s^2\, ds\right)\right] \tag{7}$$

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

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

$$V(\hat{\mu}_0, \Omega_0) = \max_{\tau} \mathbb{E}\left[\int_0^{\tau} -e^{-rs}\hat{\mu}_s^2 ds + e^{-r\tau}\Big(-\bar{\theta} + \max_{\mu'} V(\mu', \Omega_\tau)\Big)\Big|\mathcal{I}_t\right] \tag{8}$$

subject to the filtering equations in Proposition 1.

We observe in Equation (8) that the estimates enter directly in the instantaneous return, while uncertainty affects the continuation value. To be more precise, uncertainty does have a negative effect on current profits that reflects the firms' permanent ignorance about the true state. However, this loss is constant and can be treated as a sunk cost; thus it is set to zero.

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

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

Since the firm has two states, the inaction region is two-dimensional, in the space of markup gap estimations and uncertainty.

Let $\bar{\mu}(\Omega)$ denote the inaction region's border as a function of uncertainty. The inaction region is described by the set:

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

The symmetry of the inaction region is inherited from the specification of the stochastic processes and the quadratic profit function. Symmetry and zero inflation imply that the optimal reset level of the markup gap is equal to $\mu' = 0$. Proposition 3 provides an analytical approximation for the inaction region's border $\bar{\mu}(\Omega)$. An evaluation of analytical approximations can be found in the Web Appendix[6].

Proposition 3 (Inaction region) *Let r and $\bar{\theta}$ small. The inaction region $\mathcal{R}$ is approximated by*

$$\bar{\mu}(\Omega_t) = \left(6\bar{\theta}\Omega_t^2\right)^{1/4} \underbrace{\mathcal{L}_{\bar{\mu}}(\Omega_t)}_{learning} \quad where \quad \mathcal{L}_{\bar{\mu}}(\Omega_t) \equiv \left[1 + \left(\frac{\Omega_t}{\Omega^*} - 1\right)\frac{24\bar{\theta}^2}{\gamma}\right]^{-1/4} \tag{9}$$

The inaction region's elasticity with respect to Ω is $\mathcal{E} \equiv \frac{1}{2} - \frac{6\bar{\theta}^2}{\gamma}\frac{\Omega}{\Omega^*}$.

Effect of uncertainty on inaction region The inaction region can be expressed as two components that depend on uncertainty. The first component captures the well-known *option value effect* (see **?** and **?**) and it is increasing in uncertainty. One of the new implications of the model with information frictions is that this option value effect becomes time-varying and driven by uncertainty. The second component, labeled as *learning effect* and denoted by $\mathcal{L}_{\bar{\mu}}(\Omega_t)$, is a factor that amplifies or dampens the option value effect depending on the ratio of current uncertainty to fundamental uncertainty $\frac{\Omega}{\Omega^*}$. When current uncertainty is high with respect to its average level $\left(\frac{\Omega}{\Omega^*} > 1\right)$, uncertainty is expected to decrease ($\mathbb{E}[d\Omega_t] < 0$) and therefore future option values also decrease. In turn, this feeds back into the current inaction region shrinking it by a factor $\mathcal{L}_{\bar{\mu}}(\Omega_t) < 1$. When uncertainty is low with respect to its average level $\left(\frac{\Omega}{\Omega^*} < 1\right)$, it is expected to increase ($\mathbb{E}[d\Omega_t] > 0$) and thus the option values in the future also increase. Analogously, this feeds back into current bands that get expanded by a factor $\mathcal{L}_{\bar{\mu}}(\Omega_t) > 1$. Finally, when $\Omega_t = \Omega^*$, the expected change in uncertainty is zero ($\mathbb{E}[d\Omega_t] = 0$) and the learning effect disappears.

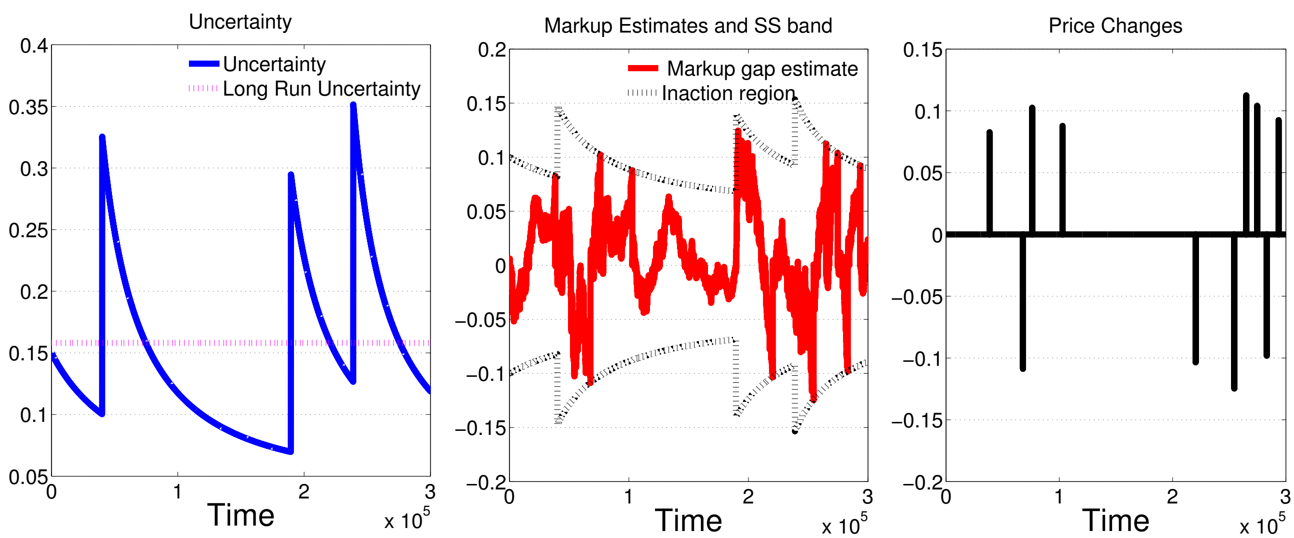
The possibility that more uncertainty could shrink the inaction region is not intuitive, as one could think that waiting a few more periods would bring a better estimate and decisions to adjust should be postponed by widening the inaction region. However, lower future uncertainty reduces the future option values. The overall effect of uncertainty on the inaction region depends on the size of the menu cost and the signal noise. Expression (9) shows that small menu costs paired with large signal noise make the learning effect close to one, implying that the elasticity of the inaction region with respect to uncertainty is close to 1/2 and thus the inaction region is increasing in uncertainty.

[6] The Web Appendix can be found at https://sites.google.com/a/nyu.edu/isaacbaley/Research.

Interplay between infrequent shocks and information frictions Without infrequent shocks, uncertainty converges to the fundamental volatility in the long run, shutting down the learning effect. The inaction region becomes constant. That is the case analyzed in the Online Appendix of **?**. As these authors show, such a model collapses to that of **?** where there is no dispersion in the size of price changes. Specifically, without infrequent shocks the steady state inaction region is constant and akin to that of a steady state model without information frictions, namely $\bar{\mu} = \left(6\bar{\theta}{\sigma_f}^2\right)^{1/4}$.

Figure II shows a particular sample path for a firm for a parametrization with small menu costs and large noise. The left panel plots the estimate of the markup gap (solid) and the inaction region (dashed). The center panel shows the evolution of uncertainty: it decreases monotonically with learning until an infrequent shock arrives and makes uncertainty jump up; then, learning brings uncertainty down again. This path is inherited by the inaction region because the calibration makes the inaction region increasing in uncertainty. The dashed horizontal line is fundamental uncertainty Ω^*. Finally, in the right panel the black vertical lines mark the times when there is a price change. These changes are triggered when the markup gap estimate touches the inaction region.

Figure II: Sample paths for one firm.



Left: Firm uncertainty (solid line) and fundamental uncertainty (horizontal dotted line). Center: Markup gap estimate (solid line) and inaction region (dotted line). Right: Price changes (vertical solid lines).

3 Implications for micro-price statistics

This section explores the implications of the price-setting model for micro price statistics.

3.1 Frequency of adjustment

How does firm uncertainty affect the frequency of adjustment? Figure II shows that after the arrival of an infrequent shock at time $t = 1$, the inaction region is wider and also the estimation is more volatile. The estimation hits the band more often and subsequent price changes are triggered. Higher uncertainty brings more price changes. The next proposition formalizes this result by characterizing analytically the expected time between price changes.

Proposition 4 (Expected time between price changes) *Let r and $\bar{\theta}$ be small. The expected time between price changes conditional on the state, denoted by $\mathbb{E}[\tau \big| \hat{\mu}_t, \Omega_t]$, is approximated as:*

$$\mathbb{E}[\tau \big| \hat{\mu}_t, \Omega_t] = \frac{\bar{\mu}(\Omega_t)^2 - \hat{\mu}_t^2}{\Omega_t^2} \underbrace{\mathcal{L}_\tau(\Omega_t)}_{learning} \quad where \quad \mathcal{L}_\tau(\Omega_t) \equiv \left[1 + \left(\frac{\Omega_t}{\Omega^*} - 1\right)\left(\frac{24\bar{\theta}^2}{\gamma}\right)\left(\frac{3\gamma + \bar{\theta}^2}{\gamma + 2\bar{\theta}^2}\right)\right] \quad (10)$$

and it is a decreasing function of uncertainty: $\frac{\partial \mathbb{E}[\tau \big| \hat{\mu}_t, \Omega_t]}{\partial \Omega_t} < 0$.

The expected time between price changes has two terms that multiply. The first term $\frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}$ is standard in these problems; it says that the closer the current markup gap is to the border of the inaction region, then the shorter the expected time for the next adjustment. This term is decreasing in uncertainty with an elasticity larger than unity in absolute value, and again the novelty is that this effect is time varying.

The second is a new term $\mathcal{L}_\tau(\Omega_t)$ that amplifies or dampens the standard effect and it arises in the learning model. As with the learning term in the inaction region, $\mathcal{L}_\tau(\Omega_t)$ is larger or smaller than one depending on the ratio of current to fundamental uncertainty. In any case, the elasticity of this term with respect to uncertainty is lower than unity; therefore, the overall effect of uncertainty on the expected time to adjustment is negative: a firm with high uncertainty is going to change the price more frequently than a firm with low uncertainty.

Proposition 5 provides an additional way to understand the negative relationship between uncertainty and frequency of adjustment.

Proposition 5 (Uncertainty and Frequency) *The following relationship between uncertainty, frequency of adjustment, and price dispersion holds:*

$$\mathbb{E}[\Omega^2] = \frac{\mathbb{V}[\Delta p]}{\mathbb{E}[\tau]} \quad (11)$$

The previous result, which is derived in Proposition 1 of **?** for the case of $\Omega_t = \sigma_f \ \forall t$, establishes a very intuitive link between uncertainty and price statistics. For a given price change dispersion, an increase in uncertainty dispersion has to lower the expected time between prices changes and vice-versa, for a fixed frequency, an increase in uncertainty dispersion has to increase the dispersion of price changes. This relationship proves to very extremely useful to back out an unobservable state, uncertainty, with observable price statistics. Furthermore, it highlights the connection between dispersion in uncertainty and dispersion in the price change distribution.

3.2 Hazard rate of price adjustment

The hazard rate of price adjustment $h(\tau)$ is defined as the probability of changing the price at date τ given that it has not been changed up to that time, this is $h(\tau) \equiv \frac{f(\tau)}{\int_\tau^\infty f(s)ds}$, where $f(s)$ is the distribution of stopping times. This probability is given by the probability of exiting the inaction region, or first passage time. Assume the last adjustment occurred at time t. Then the hazard rate is a function of two statistics:

i) the variance of the estimate at date τ from a time t perspective, or unconditional estimation variance, denoted as $p(\tau|\Omega_t)$:

$$\hat{\mu}_\tau | \mathcal{I}_t \sim \mathcal{N}(0, p(\tau|\Omega_t))$$

ii) the expected path of the inaction region given the information available at time t.

An analytical characterization of the hazard rate is provided in the following proposition. The presence of infrequent shocks only changes the level of the hazard rate but not its slope, thus we characterize the hazard rate assuming no infrequent shocks ($\lambda = 0$). Furthermore, the inaction region is assumed to be constant. This is also a valid assumption since what matters for the hazard rate is the size of the inaction region relative to the volatility of the uncontrolled process. The validity of both assumptions is explored in the Web Appendix where we compute the numerical hazard rate. Without loss of generality, set $t = 0$.

Proposition 6 (Hazard rate of price adjustment) *Let Ω_0 be the current level of uncertainty. Assume that there are no infrequent shocks ($\lambda = 0$) and constant inaction region $\bar{\mu}(\Omega_\tau) = \bar{\mu}_0$.*

- *Then the unconditional forecast variance $p(\tau|\Omega_0)$ is given by:*

$$p(\tau|\Omega_0) = \sigma_f^2 \tau + \gamma \underbrace{\mathcal{L}_p(\tau|\Omega_0)}_{learning} \tag{12}$$

where $\mathcal{L}_p(\tau|\Omega_0)$ is an increasing and concave function, $\mathcal{L}_p(0|\Omega_0) = 0$, $\lim_{\tau\to\infty} \mathcal{L}_p(\tau, \Omega_0) = \Omega_t$ and it is given by:

$$\mathcal{L}_p(\tau, \Omega_0) \equiv \Omega_0 - \sigma_f coth\left(\frac{1}{2}\log\left(\frac{\Omega_0 + \sigma_f}{\Omega_0 - \sigma_f}\right) + \frac{\sigma_f}{\gamma}\tau\right)$$

- *The hazard of adjusting the price at date τ is characterized by:*

$$h(\tau|\Omega_t) = \frac{\pi^2}{8}\frac{p'(\tau|\Omega_t)}{\bar{\mu}_0^2}\Psi\left(\frac{p(\tau|\Omega_0)}{\bar{\mu}_0^2}\right) \tag{13}$$

where Ψ is an increasing function, $\Psi(0) = 0$, $\lim_{x\to\infty}\Psi(x) = 1$, and it is given by:

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

- *If σ_f is small, then for τ sufficiently large, the hazard rate is decreasing.*

$$\frac{\partial h(\tau|\Omega_0)}{\partial\tau}\frac{1}{h(\tau|\Omega_0)} \approx -\frac{2}{\tau + \frac{\gamma}{\Omega_0}} < 0$$

Unconditional forecast variance $p(\tau|\Omega_0)$ in (12) captures the evolution of uncertainty. Its first part, $\sigma_f^2\tau$, refers to the linear time trend that comes from the markup estimates following a Brownian Motion. The second part, $\mathcal{L}_p(\tau|\Omega_0)$, is a additional source of variance coming from learning. Because $p(\tau|\Omega)$ accumulates the variance from all shocks received since the beginning, its first derivative reflects uncertainty (conditional forecast variance at time τ) and its second derivative reflects uncertainty growth. Concavity of $\mathcal{L}_p(\tau|\Omega)$ is a consequence of uncertainty being expected to decrease over time. This in turn generates a decreasing hazard rate.

Decreasing hazard rate The economics behind the decreasing hazard rate are as follows. Because of learning, firm uncertainty decreases with time and the weight given to new observations in the forecasting process decreases too. Since the volatility of the markup gap estimates is reduced, the probability of adjusting also decreases. A firms expects to transition from high uncertainty and frequent adjustment to low uncertainty and infrequent adjustments, generating a decreasing hazard rate. When frequent volatility σ_f is small, the slope of the hazard rate is driven by the noise volatility γ. Therefore γ can be chosen to match the shape of the hazard rate.

4 General Equilibrium Model

How does an economy with information frictions and nominal rigidities respond to aggregate nominal shocks? This section studies the response of output to a monetary shock in a standard general equilibrium framework with monopolistic firms that face the pricing-setting problem with menu costs and information frictions studied in the previous sections. The main result is that imperfect information amplifies nominal rigidities in normal times, while it reduces the effect of nominal rigidities if uncertainty is sufficiently high.

4.1 General Equilibrium Model

Environment There is a representative consumer and a continuum of monopolistic firms indexed by $z \in [0,1]$. All the shocks in the model are idiosyncratic.

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

$$\mathbb{E}\left[\int_0^\infty e^{-rt}\Big(\log C_t - N_t + \log \frac{M_t}{P_t}\Big)dt\right] \tag{14}$$

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

$$C_t = \Big(\int_0^1 \Big(A_t(z)c_t(z)\Big)^{\frac{\eta-1}{\eta}} dz\Big)^{\frac{\eta}{\eta-1}} \tag{15}$$

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

$$P_t \equiv \left[\int_0^1 \left(\frac{p_t(z)}{A_t(z)}\right)^{1-\eta} dz\right]^{\frac{1}{1-\eta}} \tag{16}$$

In the consumption bundle and the price index, $A_t(z)$ reflects the quality of the good, with higher quality providing larger marginal utility of consumption but at a higher price. These shocks are firm specific and will be described fully in the firm's problem.

The household has access to complete financial markets. The budget includes income from wages W_t, profits Π_t from the ownership of all firms, and the opportunity cost of holding cash $R_t M_t$, where R_t is the nominal interest rate.

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

$$\mathbb{E}\left[\int_0^\infty Q_t\left(P_tC_t + R_tM_t - W_tN_t - \Pi_t\right)dt\right] \leq M_0 \tag{17}$$

The household problem is to choose consumption of the different goods, labor supply and money holdings to maximize preferences (14) subject to (15), (16) and (17).

Monopolistic Firms A continuum of firms produce and sell their products in a monopolistically competitive market. They own a linear technology that uses labor as its only input: producing $y_t(z)$ units of good z requires $l_t(z) = y_t(z)A_t(z)$ units of labor, so that the marginal nominal cost is $A_t(z)W_t$ (higher quality $A_t(z)$ requires more labor input). The assumption that the quality shock enters both the production function and the marginal utility is done for tractability as it helps to condense the numbers of states of the firm into one, the markup, as in **?**. Each firm sets a nominal price $p_t(z)$ and satisfies all demand at this posted price. Given the current price $p_t(z)$, the consumer's demand $c_t(z)$, and current quality $A_t(z)$, instantaneous nominal profits of firm z equal the difference between nominal revenues and nominal costs:

$$\Pi(p_t(z), A_t(z)) = c_t(p_t(z), A_t(z))\Big(p_t(z) - A_t(z)W_t\Big)$$

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

$$V(p_0(z), A_0(z)) = \max_{\{p_{\tau_i}(z),\tau_i(z)\}_{i=1}^\infty} \mathbb{E}\left[\sum_{i=0}^\infty Q_{\tau_{i+1}(z)}\left(-\theta + \int_{\tau_i(z)}^{\tau_{i+1}(z)} \frac{Q_s}{Q_{\tau_{i+1}(z)}}\Pi(p_{\tau_i(z)}, A_s(z))ds\right)\right] \tag{18}$$

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

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

$$da_t(z) = \sigma_f W_t(z) + \sigma_u u_t(z) dq_t(z)$$

where $W_t(z)$ is a standard Brownian motion and $q_t(z)$ a Poisson counting process with arrival rate λ and Normal innovations $u_t(z) \sim \mathcal{N}(0, 1)$ as in the previous sections. Parameters $\{\sigma_f, \sigma_u, \lambda, \gamma\}$ are identical across firms. These shocks influence the optimal prices and generate the cross-sectional dispersion in price changes.

As before, firms do not observe the shocks to their quality directly. They do not learn it from observing their wage bill or revenues either. The only sources of information are noisy signals about quality:

$$ds_t(z) = a_t(z)dt + \gamma dZ_t(z)$$

where $Z_t(z)$ is an independent Brownian motion for each firm z and γ is signal noise. Each information set is $\mathcal{I}_t(z) = \sigma\{s_r(z), q_r(z); r \leq t\}$.

Money supply Money supply is constant at a level $\bar{M}$.

Equilibrium An equilibrium is a set of stochastic processes for (i) consumption strategies $c_t(z)$, labor supply N_t, and money holdings M_t for the household, (ii) pricing functions $p_t(z)$, and (iii) prices W_t, R_t, Q_t, P_t such that the household and firms optimize and markets clear at each date.

4.2 Characterization of Steady State Equilibrium

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

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

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

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

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

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

where the constants are $C \equiv K\eta^{-\eta}(\eta-1)^{\eta-1}$ and $B \equiv \frac{1}{2}K\frac{(\eta-1)^\eta}{\eta^{\eta-1}}$. The constant C does not affect the decisions of the firm and it is omitted for the calculations of decision rules; the constant B captures the curvature of the original profit function. This quadratic problem is the same as 7.

Markup gap estimation and uncertainty The markup gap is equal to

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

When the price is kept fixed (inside the inaction region), the markup gap is driven completely by the productivity process: $d\mu_t(z) = -da_t(z)$. When there is a price adjustment, the markup process is reset to its new optimal value and then it will again follow the productivity process. By symmetry of the Brownian motion without drift and the mean zero innovations of the Poisson process, $da_t(z) = -da_t(z)$. The filtering equations are as in Proposition 1, but each process is indexed by z and is independent across firms.

$$\begin{aligned} d\hat{\mu}_t(z) &= \Omega_t(z) dX_t(z) \\ d\Omega_t(z) &= \frac{\sigma_f^2 - \Omega_t^2(z)}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dq_t(z) \end{aligned}$$

where $X_t(z)$ is a standard Brownian motion for every z.

4.3 Data and calibration

The model is solved numerically as a discrete time version of the continuous time model described above. The calibration of the model is done using price statistics reported in **?**, who use BLS monthly data for a sample that is representative of consumer and producer prices except services, controlling for heterogeneity and sales. The sample is restricted to regular price changes, that is, with sales filtered out. The moments considered are the mean price change $\mathbb{E}[|\Delta p|] = 0.11$, the expected time between price changes $\mathbb{E}[\tau] = 8$ months, as well as the hazard rate with negative slope. These moments are also consistent with Dominick's database reported in **?**.

The frequency considered is weekly and the discount factor is set to $\beta = 0.96^{1/52}$ to match an annual interest risk free rate of 4%. The normalized menu cost $\bar{\theta}$ is set to 0.064 so that the expected menu cost payments $(\frac{1}{\mathbb{E}[\tau]}\theta)$ represent 0.5% of the average revenue given by $(\frac{\eta}{\eta-1})^{1-\eta}$ as in **?** and **?**. The CES elasticity of substitution between goods is set to $\eta = 6$ in order to match an average markup of 20%.

The stochastic processes are calibrated to match the price statistics. In addition to the model with information frictions, two other versions are calibrated to serve as benchmarks. These alternative models assume perfect information ($\gamma = 0$). The first model shuts down the infrequent shocks ($\lambda = 0$) and its only parameter σ_f is set to match the frequency of adjustment. The second model shuts down the frequent shocks ($\sigma_f = 0$), and its two parameters λ and σ_u match the frequency and the dispersion of price changes. The model with information frictions has an additional pa-

Table I:
Model parameters and targets

Target	Data	Model		
		Perfect Info		Info Frictions
	Monthly BLS	No infreq shocks (Golosov & Lucas)	Infreq shocks (Gertler & Leahy)	
$\mathbb{E}[\tau]$	2.3 quarters	$\sigma_f = 0.0195$	$\lambda = 0.04$	$\lambda = 0.027$
$std[\lvert\Delta p\rvert]$	0.08		$\sigma_u = 0.145$	$\sigma_u = 0.2$
$min\lvert\Delta p\rvert$	≈ 0		$\sigma_f = 0$	$\sigma_f = 0.0005$
$h(\tau)$	$slope < 0$			$\gamma = 0.15$

Monthly data from BLS in **?** converted to weekly frequency. For the slope of the hazard rate $h(\tau)$ see Figure III.

rameter to calibrate, the signal noise, which is set to $\gamma = 0.15$ to match the shape of the hazard rate[7].

Price statistics The left panel in Figure III shows the ergodic distribution of price changes for different parametrizations of the model and the data. Its symmetry comes from the assumption of zero inflation and the stochastic process. The model with only small frequent shocks [dotted line] generates a price distribution concentrated at the borders of the inaction region. The models with infrequent shocks, with and without information frictions, are able to better match the distribution of price changes with fat tails and larger dispersion.

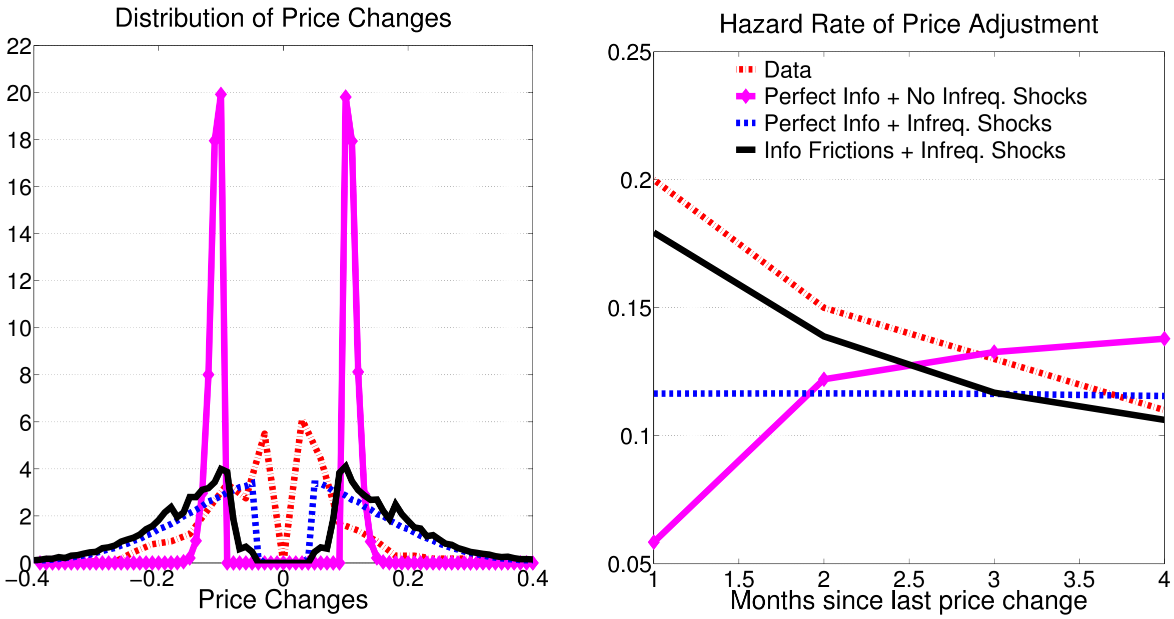
The model has some difficulty in matching small price changes because the minimum is bounded by the size of the menu cost. However, as noted by **?**, small price changes might be the result of measurement errors and not a reason to dismiss a menu cost model. In any case, economies of scope through multi-product firms can be added to generate such small price changes, see **?** and **?**. The volatility of the frequent shocks, σ_f, is set very close to zero so that the minimum level of uncertainty is also close to zero and tiny price changes may arise from certain firms.

The right panel of figure III plots the hazard rate of price adjustment. Infrequent shocks with perfect information generate a non-decreasing, almost flat hazard. Perfect information and only frequent shocks feature an increasing hazard rate. For the imperfect information model, the decreasing hazard is obtained by calibrating the signal noise γ.

All models generate the same expected time between price changes. With imperfect information and infrequent shocks the hazard rate is decreasing: after a price change, there are correlated price changes triggered by the first one. These additional price changes allow us to lower the arrival rate of Poisson shocks λ while keeping the expected duration of prices unchanged.

[7] The Web Appendix shows how the slope of the unconditional hazard rate varies with different choices of γ without changing the price change distribution.

Figure III:
Distribution of price changes and hazard rate of price adjustments for models and data.



Three models and the data from **?** and **?**.

5 Propagation of nominal shocks

This section studies the effect of unanticipated aggregate nominal shock on output.

Unanticipated permanent nominal shocks Aggregate nominal shocks are implemented as an unforeseen and permanent increase in nominal wages, which translates into a permanent decrease in markups to all firms of size δ, where δ is small. This shock is unanticipated as firms assign zero probability to it. Output responses are computed using steady state policies following **?**, Proposition 7, which shows that general equilibrium effects on policies are irrelevant.

Start with the steady state invariant distribution of firms' price changes that correspond to an initial level of money supply equal to $\bar{M}$. Then at time $t = 0$, there is an unanticipated permanent increase in the level of money supply by $\delta\%$, or in log terms:

$$\log M_t = \log \bar{M} + \delta, \quad t \geq 0$$

Since wages are proportional to the money supply, the shock translates directly into a wage increase.

Price response The ideal price index in (16) can be written in terms of the markup gaps, just multiply and divide by the nominal wages and use the definition of markups and markup gaps:

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

Then take the log difference from steady state, a first order approximation to the integral, and substitute the wage deviation for δ.

$$\ln \left(\frac{P_t}{\overline{P}} \right) \approx \delta + \int_0^1 \mu_t(z) dz$$

Output response In order to compute the output response, the equilibrium condition is that output is equal to the real wage, which in log deviations from steady state reads:

$$\ln \left(\frac{Y_t}{\overline{Y}} \right) = \delta - \ln \left(\frac{P_t}{\overline{P}} \right)$$

The cumulative effect of the monetary shock (the area under the impulse response function or the excess output above steady state) is denoted as $\mathcal{M}$ and it equals $\mathcal{M} \equiv \int_0^\infty \ln \left(\frac{Y_t}{\overline{Y}} \right) dt$. In a frictionless world, all firms would increase their price in δ to reflect the higher marginal costs and the shock would have no output effects, i.e. $\mathcal{M} = 0$. As long as the price level fails to reflect the shock there are real effects on output.

5.1 Impulse-response of output

A nominal shocks of size $\delta = 1\%$ shocks all firms. The first column of Table III computes the total output effects $\mathcal{M}$ and half-life relative to the case perfect information with no infrequent shocks ($\lambda = 0$). For this case, $\mathcal{M} = 1.41\%$ and the half-life is 1.25 months. The results in columns 2-4 are multiples of the values in the first column.

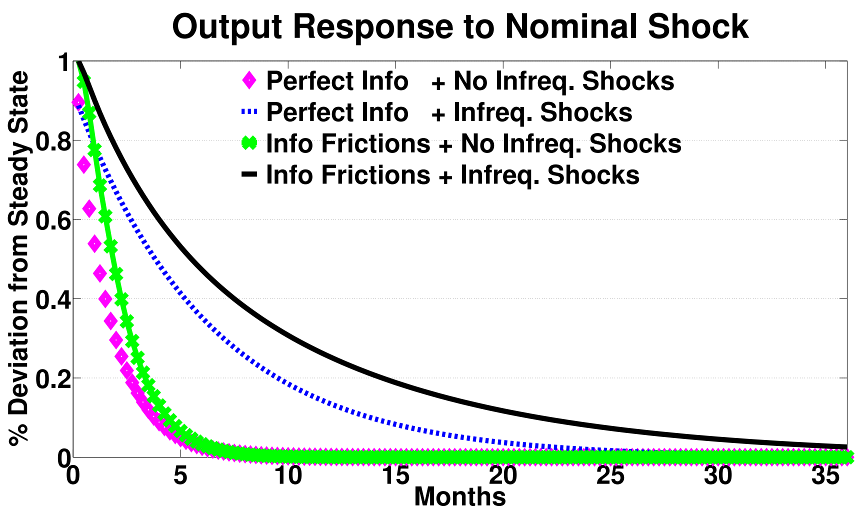
Table II: Output response to monetary shocks (as a multiple of the first column).

	Perfect Info		Info Frictions	
	(1)	(2)	(3)	(4)
Output Effect	No infreq shocks	Infreq shocks	Homog Ω	Heterog Ω
Total effect ($\mathcal{M}$)	1.0	3.2	1.4	4.6
Half-life ($t_{0.5}$)	1.0	3.9	1.2	6.1

With perfect information and $\lambda = 0$ (first column) the total output effects are $\mathcal{M} = 1.41\%$ and half-life of 1.25 months.

Figure (IV) compares the impulse-response functions of output for different information and shock structures.

Figure IV: Output response to a positive monetary shock of size $\delta = 1\%$.



Decomposing output response: selection vs. learning effects Since **?** it has been known that what matters for the flexibility of the aggregate price level in menu cost models is the presence or absence of selection effect: the firms that adjust prices are those further away from the optimal reset price and thus they have large price changes. This selection effect in turn depends on the type of idiosyncratic shocks that firms face, as subsequent work showed.

First consider the models with perfect information. The model with only small frequent shocks which is the benchmark case (Column 1 of Table III) features a large selection effect and output response is small very short lived. This is the main result in **?**. Adding large infrequent shocks (Column 2 of Table III) increases non-neutrality by breaking the selection effect, as in **?** and **?**.

Now consider the models with information frictions. The output effects not only depend on the magnitude of the selection effect but also on the speed at which firms incorporate the shock into their estimations. If information frictions are important, then the cross-sectional average of forecast errors has a large and persistent response to nominal shocks that in turns makes the response of the aggregate price level equally sluggish. In the absence of infrequent shocks (Column 3 of Table III), the steady state features no heterogeneity in uncertainty and it is equal to fundamental uncertainty, which in this case is equal to the frequent volatility for all firms $\Omega_t(z) = \Omega^* = \sigma_f$. Homogeneous uncertainty implies homogenous forecast errors.

The full model with information frictions $\gamma > 0$, infrequent shocks $\lambda > 0$ and small σ_f (Column 4 of Table III) is able to generate six times as much persistence – measured through the half life of output response– than the benchmark case. Firms' fundamental uncertainty Ω^* is close to zero and the bulk of uncertainty movements come from the arrival of infrequent shocks. Since receiving an infrequent shock implies that forecasts fully incorporate the shock (the weight on new information is close to one) and there is a price adjustment for the same amount (the shocks are large), price adjustment mostly depends on the arrival rate. This positive and high correlation between updating information sets and updating prices resembles the model by **?**. This high correlation generates similar aggregate dynamics as their sticky-information model in which this correlation is exogenously set to one. In this case, forecast errors drive almost all the action in the output effects.

5.2 Aggregate uncertainty and nominal shocks

This section explores the output response to a nominal shock when it is interacted with a synchronized uncertainty shock across all firms. The exercise is motivated by recent literature that explores the interaction of aggregate uncertainty and monetary policy. A first example is **?** that estimates the interaction effects of different measures of economic uncertainty (VIX, policy uncertainty index of **?**, macro and micro uncertainty measures in **?**, among other measures) with monetary policy shocks as identified through structural vector autoregressions. It shows that output responses to a monetary policy shock are at least halved when uncertainty measures are in their upper compared to their lower deciles. Two other examples use the dispersion of the price change distribution as their measure of aggregate uncertainty. **?** estimates a time-varying New Keynesian Phillips Curve and finds a steeper slope during times of high price change dispersion, and **?** documents a positive relationship between exchange-rate pass-through and price change dispersion.

Through the lens of the model, nominal shocks become neutral when uncertainty is high for all firms. To show this, the nominal shock is interacted with an exogenous correlated shock of uncertainty across firms. Besides the nominal shock of size $\delta = 1\%$, there is a once and for all increase in firms' uncertainty $\Omega(z)$ in $\kappa\Omega^*$ units , where κ takes different values across experiments $\kappa \in \{0, 1, 4\}$. Figure V shows the results from the uncertainty experiments. A monetary shock paired with an uncertainty shock of one times its steady state level ($\kappa = 1$) cuts output response by half, and if the uncertainty shock is four times its steady state level ($\kappa = 4$), the output effects are almost muted. These results come from the effects of uncertainty in the frequency of adjustment.

Figure V: Responses to positive monetary shock and correlated uncertainty shocks.

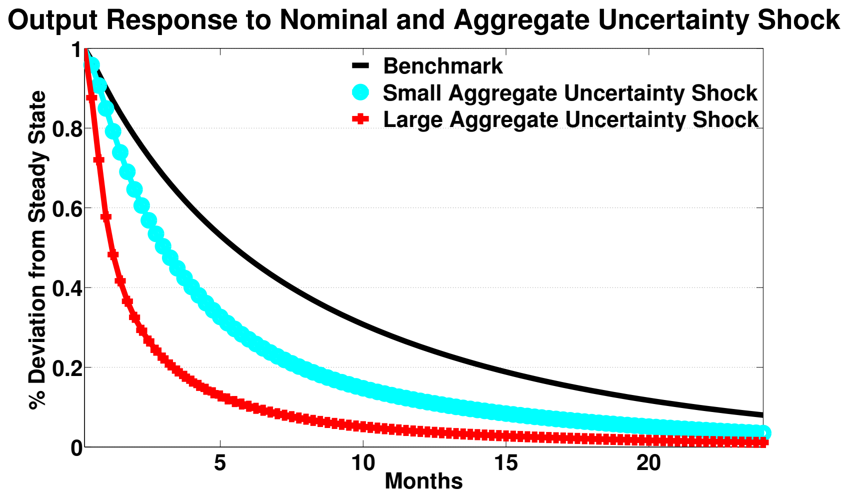


Table III: Output response to a monetary shocks paired with aggregate uncertainty shock

Output Effect	No Ω shock $\kappa = 0$	Small Ω shock $\kappa = 1$	Large Ω shock $\kappa = 4$
	Aggregate Uncertainty Shock		
Total effect $(\mathcal{M})$	4.6	2.6	1.0
Half-life $(t_{0.5})$	6.1	3.8	1.8

As multiple of benchmark case, column (1) in Table II above.

6 Conclusions

Central banks around the world use models that produce large and persistent output responses to monetary shocks at business cycle frequency. These models have two main building blocks, namely Calvo pricing and strategic complementarities, which together generate the desired inertia in inflation. However, the evidence on such mechanisms is controversial. The cross-sectional heterogeneity in price-setting observed in micro data rejects the constant probability of adjustment of Calvo pricing, and the assumption of complementarities makes reset inflation (inflation conditional on price adjusters) very persistent, while in the data it shows no persistence, as pointed out by **?**. The model developed in this paper is an alternative to the assumption of complementarities that generates persistent output responses, while also explaining new micro evidence on time varying frequencies at the individual level as well as macro evidence on monetary policy during uncertain times.

For now the paper remains silent about the source of the infrequent shocks. If these shocks come from a source within the firm, for instance an innovation process, then the decision to create a new product will be linked to its pricing decision and the pricing of other products in the firm. If these shocks come from an aggregate source, such as a tax change, then the correlation of these shocks across firms must be incorporated into the model.

A Appendix (PRELIMINARY)

Proposition 1 (Filtering equations) *Let the following processes define the markup gap and the signal*

$$\begin{aligned}
(\textit{state}) \qquad d\mu_t &= \sigma_f dW_t + \sigma_u u_t dq_t \\
(\textit{observation}) \qquad ds_t &= \mu_t dt + \gamma dZ_t \\
(\textit{initial conditions}) \qquad \mu_0 &= s_0 = 0 \\
\textit{where} \quad W_t, Z_t \sim \textit{Wiener Process}, \quad q_t &\sim \textit{Poisson}(\lambda), \quad u_t \sim \mathcal{N}(0,1)
\end{aligned}$$

Let the information set be given by $I_t = \sigma\{s_r, q_r; r \leq t\}$. Denote the best estimate (in an MSE sense) of the markup gap as $\hat{\mu}_t \equiv \mathbb{E}[\mu_t|\mathcal{I}_t]$ and the variance of estimate as $\Sigma_t \equiv \mathbb{V}[\mu_t|\mathcal{I}_t] = \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|\mathcal{I}_t]$. Finally define belief uncertainty as $\Omega_t \equiv \frac{\Sigma_t}{\gamma}$. Then these processes evolve according to:

$$\begin{aligned}
(\textit{estimate}) \qquad d\hat{\mu}_t &= \Omega_t dX_t \\
(\textit{uncertainty}) \qquad d\Omega_t &= \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dq_t \\
(\textit{innovation}) \qquad dX_t &= \frac{1}{\gamma}\left(ds_t - d\hat{\mu}_t\right) = \frac{1}{\gamma}(\mu_t - \hat{\mu}_t)dt + dZ_t
\end{aligned}$$

with initial conditions $(\hat{\mu}_0, \Omega_0, X_0)$.

Proof. Step 1 of the proof shows that given the information set, the state and signals at each date t are Gaussian random variables; furthermore, the signals have continuous paths. Step 2 applies the Kalman-Bucy filter to the system in the absence of infrequent shocks. Step 3 extends the formulas to include the arrival of infrequent shocks.

Step 1 Write the markup gap at time t as:

$$\mu_t = \sigma_f W_t + \sigma_u \sum_{k=0}^{q_t} u_k$$

where q_t is the total number of infrequent shocks received up to t. Since q_t is in the information set, the markup gap is the sum of a Wiener process and the sum of q_t Gaussian innovations u_k; therefore the random variable $\mu_t|\{q_r, r \leq t\}$ is Gaussian as well. Note that $\mathbb{E}[\mu_t|\mathcal{I}_t] = 0$ and $\mathbb{V}[\mu_t|\mathcal{I}_t] = (\sigma_f^2 + \lambda\sigma_u^2)t$. Now rewrite the signal as

$$s_t = \int_0^t \mu_s ds + \gamma Z_t$$

which is the sum of an integral of Gaussian variables and a Wiener process, and therefore it is continuous and Gaussian.

Step 2 Apply the Kalman-Bucy filtering equations to the state-signal system assuming there are no infrequent shocks ($\lambda = 0$) to obtain:

$$\begin{aligned}
d\hat{\mu}_t &= \frac{\Sigma_t}{\gamma^2}\left(ds_t - \hat{\mu}_t dt\right) \\
d\Sigma_t &= \left(\sigma_f^2 - \frac{\Sigma_t^2}{\gamma^2}\right) dt
\end{aligned}$$

- For the estimate equation, substitute the expression for the signal ds_t to get:

$$d\hat{\mu}_t = \frac{\Sigma_t}{\gamma}\left[\frac{1}{\gamma}(\mu_t - \hat{\mu}_t)dt + dZ_t\right] = \frac{\Sigma_t}{\gamma} dX_t = \Omega_t dX_t$$

where the innovation process X_t is defined as the difference between the signal and the estimate (which equals the difference between the realization of the markup gap and its estimate plus the measurement noise):

$$dX_t \equiv \frac{1}{\gamma}(\mu_t - \hat{\mu}_t)dt + dZ_t \quad \text{or analogously} \quad X_t \equiv \int_0^t \frac{1}{\gamma}(\mu_s - \hat{\mu}_s)ds + Z_t$$

From Lemma 6.2.6 in **?**, we have that X_t is a one-dimensional standard Brownian motion given the information set of the firm $\mathcal{I}_t$.

- To obtain the expression for belief uncertainty, divide both sides by signal noise γ and apply the normalization $\Omega_t = \frac{\Sigma_t}{\gamma}$ and obtain:

$$\frac{d\Sigma_t}{\gamma} = \frac{1}{\gamma}\left(\sigma_f^2 - \frac{\Sigma_t^2}{\gamma^2}\right)dt \quad \Longrightarrow \quad d\Omega_t = \frac{1}{\gamma}\left(\sigma_f^2 - \Omega_t^2\right)dt$$

Step 3 Now infrequent shocks are introduced. The information set is thus $\mathcal{I}_t = \{d_1\}$Let $\Delta \approx 0$ be a small period of time and $\varepsilon_t, \eta_t \sim \mathcal{N}(0,1)$ two Gaussian innovations.

- Upon the arrival of a Poisson shock, the state changes discretely as:

$$\mu_t - \mu_{t-\Delta} = \sigma_f \sqrt{\Delta}\varepsilon_t + \sigma_u u_t$$

The slope of the signals jumps, but the changes in levels is continuous:

$$s_t - s_{t-\Delta} \quad = \quad \mu_t\Delta + \gamma\sqrt{\Delta}\eta_t$$

Therefore, the innovation is given by:

$$\begin{aligned} X_t - X_{t-\Delta} &= \frac{1}{\gamma}(\mu_t - \hat{\mu}_t)\Delta + \sqrt{\Delta}\eta_t = \frac{1}{\gamma}(\mu_t - \mu_{t-\Delta})\Delta + \frac{1}{\gamma}(\mu_{t-\Delta} - \hat{\mu}_t)\Delta + \sqrt{\Delta}\eta_t \\ &= \frac{1}{\gamma}(\sigma_f\sqrt{\Delta}\varepsilon_t + \sigma_u u_t)\Delta + \frac{1}{\gamma}(\mu_{t-\Delta} - \hat{\mu}_t)\Delta + \sqrt{\Delta}\eta_t \end{aligned}$$

but taking the limit $\Delta \to 0$ shows that the innovation process is continuous. This also makes the estimation $\hat{\mu}_t$ a continuous variable: $\hat{\mu}_{t-\Delta} = \hat{\mu}_t$ on the set $\{dq_t = 1\}$. Therefore, regardless of the arrival of an infrequent shocks, the estimates are given by: $d\hat{\mu}_t = \Omega_t dX_t$.

- Regarding the evolution of the variance when a Poisson shock arrives, it can be approximated in a small interval of time Δ as:

$$\begin{aligned} \mathbb{V}[\hat{\mu}_t|\mathcal{I}_t] = \mathbb{E}[(\mu_t - \hat{\mu}_t)^2|\mathcal{I}_t] &= \mathbb{E}[(\mu_{t-\Delta} + \sigma_f\sqrt{\Delta}\varepsilon_t + \sigma_u u_t - \hat{\mu}_t)^2|\mathcal{I}_t] \\ &= \mathbb{E}[(\mu_{t-\Delta} - \hat{\mu}_{t-\Delta})^2|\mathcal{I}_t] + \sigma_f^2\Delta + \sigma_u^2 \\ &= \mathbb{V}[\hat{\mu}_{t-\Delta}|\mathcal{I}_t] + \sigma_f^2\Delta + \sigma_u^2 \end{aligned}$$

where in the second line we have used the continuity of the estimate $\hat{\mu}_t = \hat{\mu}_{t-\Delta}$ and that shocks are uncorrelated. Divide by γ to obtain:

$$\frac{\mathbb{V}[\hat{\mu}_t|\mathcal{I}_t] - \mathbb{V}[\hat{\mu}_{t-\Delta}|\mathcal{I}_t]}{\gamma} = \frac{\sigma_f^2}{\gamma}\Delta + \frac{\sigma_u^2}{\gamma}$$

which by definition of Ω_t is equal to:

$$\Omega_t - \Omega_{t-\Delta} = \frac{\sigma_f^2}{\gamma}\Delta + \frac{\sigma_u^2}{\gamma}$$

Taking $\Delta \to 0$, we see that belief uncertainty has a jump when the Poisson shock arrives:

$$\Omega_t - \Omega_{t-\Delta} = \frac{\sigma_u^2}{\gamma}$$

■

Lemma 1 (Long run belief uncertainty) *If* $\Omega_t = \Omega^* \equiv \sqrt{\sigma_f^2 + \lambda\sigma_u^2}$, *then* $\mathbb{E}\left[\frac{d\Omega_t}{dt}\Big|\mathcal{I}_t\right] = 0$.

Proof. Let Ω^* be such that the expected change in uncertainty is equal to zero, this is:

$$0 \quad = \quad \mathbb{E}\left[\frac{d\Omega_t}{dt}\Big|\mathcal{I}_t\right] = \mathbb{E}\left[\sigma_f^2 - \Omega_t^2 + \sigma_u^2 \frac{dq_t}{dt}\right]$$

The solution yields:

$$\Omega_t^2 = \sigma_f^2 + \lambda\sigma_u^2 \qquad \Longrightarrow \Omega_t^* = \sqrt{\sigma_f^2 + \lambda\sigma_u^2}$$

If $\Omega_t < \Omega^*$, then $\mathbb{E}\left[\frac{d\Omega_t}{dt}\Big|\mathcal{I}_t\right] > 0$ so uncertainty is expected to increase. The opposite holds for $\Omega_t > \Omega^*$. ■

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

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

subject to the filtering equations in Lemma 1.

Proof.

- Using the definition of variance, we can write:

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

- 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:

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

- Use the Law of Iterated Expectations to take expectation of the profit function given the information set at time s. Use the decomposition above to write the problem in terms of estimates:

$$\mathbb{E}\left[\sum_{i=0}^{\infty} e^{-r\tau_{i+1}} \left(-\theta - \int_{\tau_i}^{\tau_{i+1}} e^{-r(s-\tau_{i+1})} B\mathbb{E}\left[\mu_s^2\Big|\mathcal{I}_s\right] ds\right)\right]$$
$$\mathbb{E}\left[\sum_{i=0}^{\infty} e^{-r\tau_{i+1}} \left(-\theta - \int_{\tau_i}^{\tau_{i+1}} e^{-r(s-\tau_{i+1})} (B\hat{\mu}_s^2 + \Omega^{*2}s) ds\right)\right]$$
$$\mathbb{E}\left[\sum_{i=0}^{\infty} e^{-r\tau_{i+1}} \left(-\theta - \int_{\tau_i}^{\tau_{i+1}} e^{-r(s-\tau_{i+1})} B\hat{\mu}_s^2 ds\right)\right] - \underbrace{\Omega^{*2}\mathbb{E}\left[\sum_{i=0}^{\infty} \int_{\tau_i}^{\tau_{i+1}} se^{-rs} ds\right]}_{sunk\ cost}$$

The term inside the expectation at the end of the previous expression is equal to:

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

where the sum is telescopic and all terms except the first cancel out. The last term then becomes:

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

This constant quantity is a sunk cost for the firm, coming from the fact that she will never learn the true realization of the markup gap. Since she cannot take any action to alter its value, so it can be ignored. Even more, we can divide all the expression by B and denote $\bar{\theta} = \frac{\theta}{B}$.

- Now consider the state $(\hat{\mu}_t, \Omega_t)$. Using the Principle of Optimality, the sequential problem can be expressed as a sequence of stopping time problems:

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

subject to the forecasting equations given above.

■

Proposition 3 (Inaction Region with Imperfect Information) *For small menu costs θ the boundary of the inaction region $\mathcal{R} = \{(\mu, \Omega) : |\mu| \leq \bar{\mu}(\Omega)\}$, can be approximated as:*

$$\bar{\mu}(\Omega_t) = \underbrace{\left(6\bar{\theta}\Omega_t^2\right)^{1/4}}_{option} \underbrace{\mathcal{L}_{\bar{\mu}}(\Omega_t)}_{learning} \quad where \quad \mathcal{L}_{\bar{\mu}}(\Omega_t) \equiv \left[1 + \frac{24\bar{\theta}^2}{\gamma}\left(\frac{\Omega_t}{\Omega^*} - 1\right)\right]^{-1/4}$$

and the new markup gap is set to zero: $x = 0$. The elasticity of the band with respect to Ω is $\mathcal{E} \equiv \frac{\partial \ln \bar{\mu}(\Omega)}{\partial \ln \Omega} = \frac{1}{2} - \frac{6\bar{\theta}^2}{\gamma}\frac{\Omega}{\Omega^*}$

Proof. The plan for the proof is as follows. First the Bellman equation inside the inaction region is established. Second, the HJB equation is derived using Itō's Lemma. A first order approximation is used to compute the value function when an infrequent shock arrives. Then the border and smooth pasting conditions are established.

1. **Notation:** Partial derivatives are denoted as

$$V_{\hat{\mu}} \equiv \frac{\partial V}{\partial \hat{\mu}}, \qquad V_{\hat{\mu}^2} \equiv \frac{\partial^2 V}{\partial \hat{\mu}^2}, \qquad V_{\Omega} \equiv \frac{\partial V}{\partial \Omega}, \qquad V_{\hat{\mu}^2,\Omega} \equiv \frac{\partial^3 V}{\partial \Omega \partial \hat{\mu}^2}$$

2. **Bellman Equation**: Let dt be a small interval of time and $1 - e^{-\lambda dt}$ be the probability of receiving an infrequent shock (that creates an uncertainty jump). Then the value of the firm inside the inaction region solves the following the Bellman equation:

$$\begin{aligned} V(\hat{\mu}_t, \Omega_t) &= -\hat{\mu}_t^2 dt + (1 - e^{-\lambda dt})e^{-rdt}\mathbb{E}\left[V\left(\hat{\mu}_{t+dt}, \Omega_{t+dt} + \frac{\sigma_u^2}{\gamma}\right)\right] + e^{-(\lambda+r)dt}\mathbb{E}\left[V(\hat{\mu}_{t+dt}, \Omega_{t+dt})\right] \\ &= -\hat{\mu}_t^2 dt + e^{-rdt}\mathbb{E}\left[V\left(\hat{\mu}_{t+dt}, \Omega_{t+dt} + \frac{\sigma_u^2}{\gamma}\right)\right] - e^{-(\lambda+r)dt}\mathbb{E}\left[V\left(\hat{\mu}_{t+dt}, \Omega_{t+dt} + \frac{\sigma_u^2}{\gamma}\right) - V(\hat{\mu}_{t+dt}, \Omega_{t+dt})\right] \end{aligned}$$

Start with a second order approximation (writing V_{t+dt} instead of $V(\hat{\mu}_t, \Omega_t)$):

$$\begin{aligned} V_{t+dt} &= V_t + V_{\hat{\mu}} d\hat{\mu}_t + V_{\Omega} d\Omega_t + \frac{1}{2}V_{\hat{\mu}^2}(d\hat{\mu}_t)^2 + \frac{1}{2}V_{\Omega^2}(d\Omega_t)^2 + V_{\hat{\mu}\Omega}(d\hat{\mu}_t)(d\Omega_t) \\ &= V_t + V_{\hat{\mu}}\Omega_t dW_t + V_{\Omega}\left[\frac{\sigma^2 - \Omega_t^2}{\gamma}dt + \frac{\sigma_u^2}{\gamma}dq_t\right] + \frac{1}{2}V_{\hat{\mu}^2}\Omega_t^2 dt + \frac{1}{2}V_{\Omega_t^2}(d\Omega_t)^2 + V_{\hat{\mu}\Omega}(d\hat{\mu}_t)(d\Omega_t) \\ &= V_t + V_{\hat{\mu}}\Omega_t dW_t + V_{\Omega}\left[\frac{\sigma^2 - \Omega_t^2}{\gamma}dt + \frac{\sigma_u^2}{\gamma}dq_t\right] + \frac{1}{2}V_{\hat{\mu}^2}\Omega_t^2 dt \end{aligned}$$

where in the second line we have substituted $d\hat{\mu}_t = \Omega_t dW_t$ and $(d\hat{\mu}_t)^2 = \Omega_t^2 dt$ and the evolution of uncertainty $d\Omega_t$; in the third line, all terms with order larger than dt are set to zero (i.e. $(d\hat{\mu}_t)(d\Omega_t) = 0 = (d\Omega_t)^2$). Applying expectations, subtracting V_t on both sides and taking the limit $dt \to 0$ obtains:

$$\mathbb{E}[dV_t] = \lim_{dt \to 0} \frac{\mathbb{E}[V_{t+dt} - V_t]}{dt} = V_{\Omega}\left[\frac{\sigma^2 - \Omega_t^2}{\gamma}\right] + \frac{1}{2}V_{\hat{\mu}^2}\Omega_t^2$$

3. **HJB:** Using Itō's Lemma, obtain the associated HJB equation:

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

Note that the odd derivatives of V with respect to $\hat{\mu}$ are all zero, and also the cross-derivatives of V across arguments. This comes from the symmetry of the value function and the separability across arguments.

4. **Approximation of jump in uncertainty**: Consider the following first order approximation on the second state:

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

Use the approximation in the second term of the HJB and grouping terms to get Ω^*:

$$rV(\hat{\mu},\Omega) \quad \approx \quad -\hat{\mu}^2 + \frac{1}{\gamma}\left[\Omega^{*2}-\Omega^2\right]V_\Omega(\hat{\mu},\Omega) + \frac{1}{2}\Omega^2 V_{\hat{\mu}^2}(\hat{\mu},\Omega)$$

The Numerical Appendix shows that the approximation is valid in the sense that the threshold computed analytically here and the exact numerical threshold coincide.

5. **Border and smooth pasting:** Observe that $\arg\max_x \; V(x,\Omega) = 0$ by symmetry over $\hat{\mu}$. Then the value matching and smooth pasting conditions are given by:

$$V(0,\Omega) - V(\bar{\mu}(\Omega),\Omega) \quad = \quad \bar{\theta} \tag{1}$$
$$V_{\hat{\mu}}(\bar{\mu}(\Omega),\Omega) \quad = \quad 0 \tag{2}$$

6. **Approximations of V, $V_{\hat{\mu}}$, $V_{\hat{\mu}^2}$ and $V_{\hat{\mu}^4}$.** For a given level of uncertainty Ω, we approximate V and $V_{\hat{\mu}}$ on their first argument with 4^{th} order Taylor approximations around the point $(0,\Omega)$ and evaluate them at $(\hat{\mu},\Omega) = (\bar{\mu}(\Omega),\Omega)$:

$$V(\bar{\mu}(\Omega),\Omega) \quad = \quad V(0,\Omega) + \frac{V_{\hat{\mu}^2}(0,\Omega)}{2!}\bar{\mu}(\Omega)^2 + \frac{V_{\hat{\mu}^4}(0,\Omega)}{4!}\bar{\mu}(\Omega)^4 \tag{3}$$
$$V_{\hat{\mu}}(\bar{\mu}(\Omega),\Omega) \quad = \quad V_{\hat{\mu}^2}(0,\Omega)\bar{\mu}(\Omega) + \frac{V_{\hat{\mu}^4}(0,\Omega)}{6}\bar{\mu}(\Omega)^3 \tag{4}$$

Note that odd terms do not appear due to the symmetry of the value function around $\hat{\mu}$. Combining equations (2) and (4) we get an expression for $V_{\hat{\mu}^2}(0,\Omega)$:

$$\frac{V_{\hat{\mu}^2}(0,\Omega)}{2} \quad = \quad -\frac{\theta}{\bar{\mu}(\Omega)^2} - \frac{V_{\hat{\mu}^4}(0,\Omega)}{24}\bar{\mu}(\Omega)^2 = -\frac{V_{\hat{\mu}^4}(0,\Omega)}{12}\bar{\mu}(\Omega)^2 \tag{5}$$

To characterize $V_{\hat{\mu}^4}$, take second derivative of HJB equation with respect to $\hat{\mu}$ and evaluate it at $\hat{\mu}=0$:

$$rV_{\hat{\mu}^2}(0,\Omega) = -2 + \frac{\Omega^2}{2}V_{\hat{\mu}^4}V(0,\Omega) + V_{\hat{\mu}^2,\Omega}(0,\Omega)\frac{[\Omega^{*2}-\Omega^2]}{\gamma} \tag{6}$$

Evaluating the limit $r \to 0$, we have that:

$$V_{\hat{\mu}^4}(0,\Omega) = \frac{4}{\Omega^2}\Big(1 + \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}\frac{\Omega^2-\Omega^{*2}}{\gamma}\Big)$$

7. **Derive threshold:** Using smooth pasting and border conditions, and equations (3) and (4) we have that

$$\bar{\theta} = \frac{V_{\hat{\mu}^4}(0,\Omega)}{24}\bar{\mu}(\Omega)^4$$

Substituting the expression for $V_{\hat{\mu}^4(0,\Omega)}$ in (6) into (3) and solving for $\bar{\mu}(\Omega)$ we obtain the expression for the cutoff value:

$$\bar{\mu}(\Omega) = (6\bar{\theta}\Omega^2)^{1/4}\Big(1 + \underbrace{\underbrace{\frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}}_{\Gamma(\Omega)} \frac{\Omega^2 - \Omega^{*2}}{\gamma}}_{\Lambda(\Omega)}\Big)^{-1/4} \tag{7}$$

Define $\Gamma(\Omega) \equiv \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}$ and $\Lambda(\Omega) \equiv \Gamma(\Omega)\frac{\Omega^2 - \Omega^{*2}}{\gamma}$, which are characterized next.

8. **Characterize** $\Gamma(\Omega)$. The cross-derivative $\Gamma(\Omega) \equiv \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2} = \frac{\partial}{\partial\Omega}\frac{V_{\hat{\mu}^2}(0,\Omega)}{2}$ is given by:

$$\begin{aligned}
\frac{\partial}{\partial\Omega}\frac{V_{\hat{\mu}^2}(0,\Omega)}{2} &= \frac{\partial}{\partial\Omega}\left[-\frac{V_{\hat{\mu}^4}(0,\Omega)\bar{\mu}(\Omega)^2}{12}\right] \\
&= \frac{\partial}{\partial\Omega}\left[-\frac{\Big(1 + \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}\frac{\Omega^2-\Omega^{*2}}{\gamma}\Big)}{3\Omega^2}\Omega\bar{\theta}^2\Big(1 + \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}\frac{\Omega^2 - \Omega^{*2}}{\gamma}\Big)^{-1/2}\right] \\
&= \frac{\partial}{\partial\Omega}\left[-\frac{12\bar{\theta}^2}{\Omega}\Big(1 + \frac{V_{\hat{\mu}^2,\Omega}(0,\Omega)}{2}\frac{\Omega^2 - \Omega^{*2}}{\gamma}\Big)^{1/2}\right]
\end{aligned}$$

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

$$\Gamma(\Omega) = \frac{\partial}{\partial\Omega}\left[-\frac{12\bar{\theta}^2}{\Omega}\Big(1 + \Gamma(\Omega)\frac{\Omega^2 - \Omega^{*2}}{\gamma}\Big)^{1/2}\right] \tag{8}$$

9. **Characterize** $\Lambda(\Omega)$: Note that: $\Lambda(\Omega^*) = 0$, $\Lambda'(\Omega^*) = 2\frac{\Omega^*}{\gamma}\Gamma(\Omega^*)$, and using (8) $\Gamma(\Omega^*) = \frac{12\bar{\theta}^2}{\Omega^{*2}}$. A first order Taylor approximation of $\Lambda(\Omega)$ around Ω^* yields:

$$\Lambda(\Omega) = \Lambda(\Omega^*) + \Lambda'(\Omega^*)(\Omega - \Omega^*) = 2\Omega^*\Gamma(\Omega^*)\frac{\Omega - \Omega^*}{\gamma} = \frac{24\bar{\theta}^2}{\gamma}\left[\frac{\Omega}{\Omega^*} - 1\right]$$

10. Finally, substitute $\Lambda(\Omega)$ into the cutoff value to obtain:

$$\bar{\mu}(\Omega) \approx (6\bar{\theta}\Omega^2)^{1/4}\Big(1 + \frac{24\bar{\theta}^2}{\gamma}\left[\frac{\Omega}{\Omega^*} - 1\right]\Big)^{-1/4} \tag{9}$$

Elasticity Now we compute the elasticity of the cutoff with respect to the forecast variance, which is given by $\mathcal{E}_{\phi,\Omega} \equiv \frac{\partial \ln\bar{\mu}(\Omega)}{\partial \ln\Omega}$. Applying logs to (9) we obtain:

$$\ln\bar{\mu}(\Omega) = \frac{1}{2}\ln\Omega - \frac{1}{4}\ln\Big(1 + \frac{24\bar{\theta}^2}{\gamma}\left[\frac{\Omega}{\Omega^*} - 1\right]\Big)$$

Since $\ln(1 + x) \approx x$ for x small, we approximate the previous expression with:

$$\ln\bar{\mu}(\Omega) \approx \frac{1}{2}\ln\Omega - \frac{6\bar{\theta}^2}{\gamma}\left[\frac{\Omega}{\Omega^*} - 1\right] = \frac{1}{2}\ln\Omega - \frac{6\bar{\theta}^2}{\gamma\Omega^*}e^{\ln\Omega} + \frac{6\bar{\theta}^2}{\gamma\Omega^*}$$

Taking the derivative we get the result:

$$\mathcal{E}_{\phi,\Omega} \equiv \frac{\partial \ln\bar{\mu}(\Omega)}{\partial \ln\Omega} = \frac{1}{2} - \frac{6\bar{\theta}^2}{\gamma}\frac{\Omega}{\Omega^*}$$

■

Proposition 4 (Expected time between price changes) *Let r and $\bar{\theta}$ be small. Then the expected time between* *price changes conditional on a current state* $(\hat{\mu}, \Omega)$ *is given by:*

$$\mathbb{E}[\tau \big| \hat{\mu}, \Omega] = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}\left[1 + \Big(\frac{\Omega}{\Omega^*} - 1\Big)C\right] \quad \textit{with constant } C = \frac{3\gamma + \bar{\theta}^2}{\gamma + 2\bar{\theta}^2}\frac{24\bar{\theta}^2}{\gamma}$$

Proof.

1. **HJB Equation and border condition** Define $T(\hat{\mu}, \Omega) = \mathbb{E}[\tau \big| \Omega, \hat{\mu}]$, then $T(\hat{\mu}, \Omega)$ satisfies

$$0 = 1 + \lambda\left[T\left(\hat{\mu}, \Omega + \frac{\sigma_u^2}{\gamma}\right) - T(\hat{\mu}, \Omega)\right] + \frac{(\sigma_f^2 - \Omega^2)}{\gamma}T_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2}T_{\hat{\mu}^2}(\hat{\mu}, \Omega) \tag{10}$$

   with the border condition given by

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

2. **Approximation of uncertainty jump:** Approximate the uncertainty jump with a linear approximation on the second state:

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

   Using this approximation and the definition of Ω^* , the HJB in (10) reads as:

$$0 = 1 + \frac{\Omega^{*2} - \Omega^2}{\gamma}T_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2}T_{\hat{\mu}^2}(\hat{\mu}, \Omega) \tag{11}$$

3. **Taylor approximation and border condition**: Approximate $T(\hat{\mu}, \Omega)$ with a second order Taylor approximation around a zero markup gap $\hat{\mu} = 0$.

$$T(\hat{\mu}, \Omega) = T(0, \Omega) + \frac{T_{\hat{\mu}^2}(0, \Omega)}{2}\hat{\mu}^2 \tag{12}$$

   - Using the border condition in (12), we get an expression for $T(0, \Omega)$

$$T(0, \Omega) = -\frac{T_{\hat{\mu}^2}(0, \Omega)}{2}\bar{\mu}(\Omega)^2 \tag{13}$$

   - From the HJB equation (11), we get an expression for $T_{\hat{\mu}^2}(0, \Omega)$:

$$\frac{T_{\hat{\mu}^2}(0, \Omega)}{2} = \frac{-1}{\Omega^2}\left[1 + T_\Omega(0, \Omega)\frac{\Omega^{*2} - \Omega^2}{\gamma}\right] \tag{14}$$

   - Substituting (13) and (14) into (12)

$$T(\hat{\mu}, \Omega) = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}\left[1 + \underbrace{T_\Omega(0, \Omega)\frac{\Omega^{*2} - \Omega^2}{\gamma}}_{\Lambda(\Omega)}\right] \tag{15}$$

   where $\Lambda(\Omega)$ measures the effect of changes in uncertainty on the expected time taking into account the distance from fundamental volatility.

4. **Approximate** $\Lambda(\Omega) \equiv T_\Omega(0, \Omega)\frac{\Omega^{*2} - \Omega^2}{\gamma}$**.** A first order Taylor approximation around Ω^* yields:

$$\Lambda(\Omega) = \Lambda(\Omega^*) + \Lambda_\Omega(\Omega^*)\frac{\Omega - \Omega^*}{\gamma} = 0 - 2\Omega^* T_\Omega(0, \Omega^*)\frac{\Omega - \Omega^*}{\gamma}$$

- To characterize $T_\Omega(0,\Omega^*)$, take the partial derivative of (15) with respect to Ω and evaluate it at $(0,\Omega^*)$:

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

  where the elasticity $\mathcal{E}_{\bar{\mu}(\Omega),\Omega} = \frac{1}{2} - \frac{6\bar{\theta}^2}{\gamma}$ and for $\frac{\phi(\Omega^2)}{\Omega} = \bar{\theta}^2$ from Proposition 4.
- Substituting back in the expression for $\Lambda(\Omega)$ we arrive to

$$\Lambda(\Omega) = -\frac{2\Omega^*}{\gamma}T_\Omega(0,\Omega^*)(\Omega-\Omega^*) = \left(1+\frac{\bar{\theta}^2}{3\gamma}\right)\left(\frac{2\bar{\theta}^2}{\gamma+2\bar{\theta}^2}\right)\left(\frac{\Omega}{\Omega^*}-1\right)$$

5. **Conclude** Finally, we arrive at the result

$$T(\hat{\mu},\Omega) = \frac{\bar{\mu}(\Omega)^2 - \hat{\mu}^2}{\Omega^2}\left[1 + \Big(1+\frac{\bar{\theta}^2}{3\gamma}\Big)\Big(\frac{2\bar{\theta}^2}{\gamma+2\bar{\theta}^2}\Big)\Big(\frac{\Omega}{\Omega^*}-1\Big)\right]$$

■

Proposition 5 (General identification of uncertainty) *The following relationship between uncertainty and price statistics always holds for any stopping time distribution:*

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

Proof. See Proposition 1 in **?** for a derivation of this result for the case of a fixed uncertainty level $\Omega_t = \sigma$. Here we extend the proof for the case of moving uncertainty. The process for markup estimation is given by $d\hat{\mu}_t = \Omega_t dB_t$. Using Ito'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. Using the Optional Sampling Theorem we have that

$$\mathbb{E}\left[\hat{\mu}_\tau^2 - \int_0^\tau \Omega_s^2 ds\Big|(\mu,\Omega) = (0,\tilde{\Omega})\right] = 0 \tag{16}$$

since the stochastic process inside the $\hat{\mu}_t^2 - \int_0^t \Omega_s^2 ds$ is a martingale. Therefore we can write (16) as:

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

Now we will integrate both sides using the renewal distribution $S(\Omega)$ with density $s(\Omega)$. This is the distribution of uncertainty of adjusting firms. The LHS is equal to the expectation of the square of price changes or the variance of price changes since the mean is zero

$$\int_0^\infty \mathbb{E}\left[\hat{\mu}_\tau^2\Big|(\mu_0,\Omega_0) = (0,\tilde{\Omega})\right] dS(\tilde{\Omega}) = \mathbb{E}[(\Delta p)^2] = \mathbb{V}[(\Delta p)] \tag{17}$$

On the RHS, note that $\int_0^\infty \mathbb{E}[\int_0^\tau \Omega_s^2 ds|(\mu_0,\Omega_0) = (0,\tilde{\Omega})]dS(\tilde{\Omega})$ is the expected local time $\mathcal{L}$ for the payoff function Ω_s^2, and therefore we can express it in the state domain instead of the time domain (see **?**):

$$\begin{aligned}
\int_0^\infty \mathbb{E}\left[\int_0^\tau \Omega_s^2 ds \Big| (\mu_0, \Omega_0) = (0, \tilde{\Omega})\right] s(\tilde{\Omega}) d\tilde{\Omega} &= \int_0^\infty \left(\int_{\hat{\mu},\Omega} \mathcal{L}(0, \tilde{\Omega}; \hat{\mu}, \Omega)\Omega^2 d\hat{\mu} d\Omega\right) s(\tilde{\Omega}) d\tilde{\Omega} \\
&= \int_{\hat{\mu},\Omega} \left(\int_0^\infty \mathcal{L}(0, \tilde{\Omega}; \hat{\mu}, \Omega) s(\tilde{\Omega}) d\tilde{\Omega}\right) \Omega^2 d\hat{\mu} d\Omega \\
&= \mathbb{E}[\tau] \int_{\hat{\mu},\Omega} \left(\int_0^\infty \frac{\mathcal{L}(0, \tilde{\Omega}; \hat{\mu}, \Omega)}{\mathbb{E}[\tau]} s(\tilde{\Omega}) d\tilde{\Omega}\right) \Omega^2 d\hat{\mu} d\Omega \\
&= \mathbb{E}[\tau] \int_{\hat{\mu},\Omega} \Omega^2 f(\hat{\mu}, \Omega) d\hat{\mu} d\Omega \\
&= \mathbb{E}[\tau]\mathbb{E}[\Omega^2] \qquad (18)
\end{aligned}$$

where f is the joint density. Putting together (17) and (18) we get the result:

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

■

Proposition 6 (Hazard rate of price adjustment) *Let $(\hat{\mu}, \Omega) = (0, \Omega_0)$ be the initial conditions. Assume $\lambda = 0$ and constant inaction region $\bar{\mu}(\Omega) = \bar{\mu}(\Omega_0)$.*

1. *The unconditional distribution of the markup gap forecast is $\hat{\mu}_t|\mathcal{I}_0 \sim \mathcal{N}(0, p(t|\Omega_0))$, where $p(t|\Omega_0)$ is the unconditional forecast variance given by:*

$$p(t|\Omega_0) = \sigma_f^2 t + \gamma \mathcal{L}_p(t|\Omega_0)$$

where $\mathcal{L}_p(t|\Omega_0)$ is increasing and concave, $\mathcal{L}_p(0|\Omega_0) = 0$ and $\lim_{t\to\infty} \mathcal{L}_p(t, \Omega_0) = \Omega_0$ and it is given by

$$\mathcal{L}_p(t, \Omega_0) \equiv \Omega_0 - \sigma_f coth\left(\frac{1}{2}\log\left(\frac{\Omega_0 + \sigma_f}{\Omega_0 - \sigma_f}\right) + \frac{\sigma_f}{\gamma}t\right)$$

2. *Then hazard rate of price adjustment is*

$$h(t|\Omega_0) = \frac{\pi^2}{8}\frac{p'(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}\Psi\left(\frac{p(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}\right)$$

where Ψ has the properties $\Psi(0) = 0$, $\Psi' > 0$ and $\lim_{x\to\infty}\Psi(x) = 1$ and it is given by

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

3. *For small σ_f and t sufficiently large, the **hazard rate is decreasing and its slope is only a function of** γ:*

$$\frac{\partial h(t|\Omega_0)}{\partial}\frac{1}{h(t|\Omega_0)} \approx -\frac{2}{t + \frac{\gamma}{\Omega_0}} < 0$$

Proof. Assume $\lambda = 0$ and initial conditions $(\hat{\mu}_0, \Omega) = (0, \Omega_0)$. Also assume that the inaction region is constant at $\bar{\mu}(\Omega) = \bar{\mu}(\Omega_0)$.

1. First we show that from time 0 perspective, the forecast evolves as: $\hat{\mu}_t|\mathcal{I}_0 \sim \mathcal{N}(0, p(t|\Omega_0))$, where $p(t|\Omega_0)$ is the unconditional forecast variance. Recall that the forecast evolves as $d\hat{\mu}_t = \Omega_t dB_t$, and because $\lambda = 0$, uncertainty evolves deterministically as $d\Omega_t = \frac{1}{\gamma}(\sigma_f - \Omega_t)$. The solution to the forecast equation with initial

condition $\hat{\mu} = 0$ is $\hat{\mu}_t = \int_0^t \Omega_s dB_s$. By definition of Ito's integral $\int_0^t \Omega_s dB_s = \lim_{(t_{i+1}-t_i)\to 0} \sum_{t_i} \Omega_{t_i}(B_{t_{i+1}} - B_{t_i})$. Given the Normal distribution of the increments of the Brownian motion B and that Ω_{t_i} is deterministic, we have that for each t_i, $\Omega_{t_i}(B_{t_{i+1}} - B_{t_i})$ has a Normal distribution too. Since the limit of Normal variables is normal, $\int_0^t \Omega_s dB_s \sim \mathcal{N}(0, p(t))$, where we defined the unconditional variance $p(t|\Omega_0) \equiv \mathbb{E}[\hat{\mu}_t^2 | \mathcal{I}_0]$.

We characterize the evolution of the unconditional variance. By orthogonality of the innovation to the forecast, $\mu_t - \hat{\mu}_t \perp \hat{\mu}_t$, we have that

$$\mathbb{E}_t[(\mu_t - \hat{\mu}_t)\mu_t] = \mathbb{E}_t[(\mu_t - \hat{\mu_t})^2] = \Sigma_t \tag{19}$$

Use the definition of unconditional variance and the result (19) we obtain:

$$p(t) \equiv \mathbb{E}[\hat{\mu}_t^2] = \mathbb{E}[(\hat{\mu}_t - \mu_t)^2] - 2\mathbb{E}[(\hat{\mu}_t - \mu_t)\mu_t] + \mathbb{E}[\mu_t^2] = \mathbb{E}[\mu_t^2] - \Sigma_t \tag{20}$$

Now we give expressions for $\mathbb{E}[\mu_t^2]$ and Σ_t. First, since the observations evolve as $d\mu_t = \sigma_f dF_t$, we have that $\mu_t = \mu_0 + \sigma_f F_t$, with $F_0 = 0$. Therefore

$$\mathbb{E}[\mu_t^2] = \mathbb{E}[(\mu_0 + \sigma_f F_t)^2] = \mathbb{E}[\mu_0^2] + 2\mathbb{E}[\mu_0 \sigma_f (F_t - F_0)] + \mathbb{E}[\sigma_f^2 (F_t - F_0)^2] = \mathbb{E}[\mu_0^2] + \sigma_f^2 t = \Sigma_0 + \sigma_f^2 t \tag{21}$$

where we use independent increments of the Brownian motion. Second, conditional forecast variance evolves as $d\Sigma_t = \left(\sigma_f^2 - \frac{\Sigma^2}{\gamma^2}\right) dt$, and assuming an initial condition such that $\Sigma_0 > \gamma\sigma_f$, we have that the solution is

$$\Sigma_t = \gamma\sigma_F \frac{K\exp(2\sigma_f t/\gamma) + 1}{K\exp(2\sigma_f t/\gamma) - 1}, \quad K \equiv \left|\frac{\sigma_f\gamma + \Sigma_0}{\sigma_f\gamma - \Sigma_0}\right| = \left|\frac{\sigma_f + \Omega_0}{\sigma_f - \Omega_0}\right| \tag{22}$$

Substituting expressions (21) and (22) into (20) and using the definition of uncertainty $\Omega_t = \gamma\Sigma_t$, we get:

$$\begin{aligned}
p(t|\Omega_0) &= \sigma_f^2 t + \gamma\left(\Omega_0 - \Omega_t\right) \\
&= \sigma_f^2 t + \gamma\left(\Omega_0 - \sigma_f \frac{K\exp(2\sigma_f t/\gamma) + 1}{K\exp(2\sigma_f t/\gamma) - 1}\right) \\
&= \sigma_f^2 t + \gamma\left(\Omega_0 - \sigma_f coth\left(\log\left(\left|\frac{\sigma_f + \Omega_0}{\sigma_f - \Omega_0}\right|^{1/2}\right) + \frac{\sigma_f}{\gamma} t\right)\right) \\
&= \sigma_f^2 t + \gamma\mathcal{L}_p(t|\Omega_0)
\end{aligned} \tag{23}$$

where we define the learning component as:

$$\mathcal{L}_p(t|\Omega_0) \equiv \Omega_0 - \sigma_f coth\left(\frac{1}{2}\log\left(\frac{\Omega_0 + \sigma_f}{\Omega_0 - \sigma_f}\right) + \frac{\sigma_f}{\gamma} t\right)$$

2. Let $F(\sigma^2 t, \bar{\mu}_f)$ be the cumulative distribution of stopping times of a Brownian motion with accumulated variance $\sigma_f^2 t$, initial condition 0 and a symmetric stopping time region $[-\bar{\mu}_f, \bar{\mu}_f]$. The accumulated variance is given by $p(t|\Omega_0)$ as above, so the distribution of stopping times is

$$F\left(p(t, \Omega_0), \bar{\mu}(\Omega_0)\right)$$

and its derivative

$$f(t) = f(p(t, \Omega_0), \bar{\mu}(\Omega_0))p'(t|\Omega_0)$$

Using the result of the hazard rate with perfect information and $\lambda = 0$ in **?**

$$h(t|\Omega_0) = \frac{\pi^2}{8\left(\bar{\mu}(\Omega_0)\right)^2}\Psi\left(\frac{p(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}\right)p'(t|\Omega_0)$$

where

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

is increasing, first convex then concave, $\Psi(0) = 0$ and $\lim_{x\to\infty}\Psi(x) = 1$.

3. Taking derivatives of the hazard rate with respect to t:

$$\frac{\partial h(t|\Omega_0)}{\partial t} = \frac{\pi^2}{8}\left[\underbrace{\frac{d\Psi(x)}{dx}\Big|_{x=\frac{p(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}}}_{\to 0}\underbrace{\left(\frac{p'(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}\right)^2}_{>0} + \underbrace{\Psi\left(\frac{p(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}\right)\frac{p''(t|\Omega_0)}{\bar{\mu}(\Omega_0)^2}}_{<0}\right]$$

We know that if $x \to \infty$ then $\Psi(x) \to 1$, which means that the first term $\frac{d\Psi(x)}{dx} \to 0$. The second term can be written as $h(t|\Omega_0)\frac{p''(t|\Omega_0)}{p'(t|\Omega_0)}$ where the ratio of derivatives is

$$\frac{p''(t,\Omega_0)}{p'(t,\Omega_0)} = \frac{\gamma\mathcal{L}_p''}{\sigma_f^2 + \gamma\mathcal{L}_p'} \approx \frac{2\sigma_f}{\gamma}\left(1 - coth^2\left(\frac{1}{2}\log\left(\frac{\Omega_0+\sigma_f}{\Omega_0-\sigma_f}\right) + \frac{\sigma_f}{\gamma}t\right)\right) < 0$$

Since $|coth(x)| \geq 1$, we have that this term is negative. Therefore, for large t we have a negative slope.

■

Proposition 7 (Forecast errors without infrequent shocks ($\lambda = 0$)) *Without infrequent shocks, the ensembles of forecast errors and markup gap estimates evolve as:*

$$\textit{(learning)} \qquad \mathcal{F}_t = \left(\frac{1}{1+\frac{\sigma_f}{\gamma}}\right)^t \delta \tag{24}$$

$$\textit{(selection)} \qquad \mathcal{G}_t = \mathcal{G}_{t-1} - \int_0^1 \Delta p_{t-1}^{\delta}(z)dz + \frac{\sigma_f}{\gamma}\mathcal{F}_t \tag{25}$$

with $\mathcal{G}_{-1} = 0$. The term with the integral is the average price change triggered by monetary policy.

Proof. For these results, we use the discrete time version of the problem (see Web Appendix).

- **Learning Effect**: From the Kalman filtering equations for firm z, the markup gap estimate is:

$$\hat{\mu}_t(z) = \frac{\sigma_f}{\sigma_f+\gamma}\mu_t(z) + \frac{\gamma}{\sigma_f+\gamma}\hat{\mu}_{t-1}(z)$$

where we have used that in a steady state without infrequent shocks the uncertainty of all firms is the same and equal to $\Omega(z) = \Omega^* = \sigma_f$. One period ahead forecast error is given by

$$f_t(z) = \mu_t(z) - \hat{\mu}_t(z) = \frac{\gamma}{\Omega^*+\gamma}(\mu_t(z) - \hat{\mu}_{t-1}(z)) = \frac{\gamma}{\Omega^*+\gamma}(f_{t-1}(z) + \mu_t(z) - \mu_{t-1}(z))$$

Integrating across the mass of firms, we obtain the ensemble of forecast errors:

$$\mathcal{F}_t = \int_0^1 f_t(z)dz = \frac{\gamma}{\sigma_f+\gamma}\int_0^1 f_{t-1}(z)dz + \underbrace{\int_0^1 [\mu_t(z)dz - \mu_{t-1}(z)]\,dz}_{=0} = \frac{\gamma}{\sigma_f+\gamma}\mathcal{F}_{t-1}$$

where the ensembles of markup gaps differences between at t and $t-1$ is zero. Iterating backwards, the previous equation we have the result with initial condition $F_0 = \delta$:

$$\mathcal{F}_t = \left(\frac{1}{1+\frac{\sigma_f}{\gamma}}\right)^t \delta$$

The half life H is determined by:

$$H = \frac{\ln 2}{\ln\left(1+\frac{\sigma_f}{\gamma}\right)}$$

and thus the ratio of fundamental to signal volatility, $\frac{\sigma_f}{\gamma}$ determines the speed of convergence towards zero.

- **Selection Effect**:

$$\begin{aligned}
\mathcal{G}_t = \int_0^1 \hat{\mu}_t(z)dz &= \int_0^1 \left[\frac{\sigma_f}{\sigma_f+\gamma}\mu_t(z) + \frac{\gamma}{\sigma_f+\gamma}\hat{\mu}_{t-1}(z)\right] dz \\
&= \frac{\sigma_f}{\sigma_f+\gamma}\int_0^1 \mu_t(z)dz + \frac{\gamma}{\sigma_f+\gamma}\int_0^1 \hat{\mu}_{t-1}(z)\, dz \\
&= \frac{\sigma_f}{\sigma_f+\gamma}\int_0^1 f_t(z)dz + \frac{\sigma_f}{\sigma_f+\gamma}\int_0^1 \hat{\mu}_t(z)dz + \frac{\gamma}{\sigma_f+\gamma}\mathcal{G}_{t-1} \\
&= \frac{\sigma_f}{\sigma_f+\gamma}\mathcal{F}_t + \frac{\sigma_f}{\sigma_f+\gamma}\int_0^1 \hat{\mu}_t(z)dz + \frac{\gamma}{\sigma_f+\gamma}\mathcal{G}_{t-1} \\
&= \frac{\sigma_f}{\sigma_f+\gamma}\mathcal{F}_{t-1} + \int_0^1 \hat{\mu}_{t-1}(z)dz - \int_0^1 \bar{\mu}(z)dz
\end{aligned}$$

■

Proposition 8 (Forecast errors with infrequent shocks ($\lambda > 0$)) *With information frictions ($\lambda > 0$) and $\sigma_f \approx 0$, ensembles of forecast errors and markup estimates evolve as:*

$$(\textit{learning}) \qquad \mathcal{F}_t = \left(1 - \frac{\lambda\sigma_u^2}{\sigma_u^2+\gamma^2}\right)^t \delta \tag{26}$$

$$(\textit{selection}) \qquad \mathcal{G}_t = e^{-\lambda}\mathcal{G}_{t-1} + (1-e^{-\lambda})0 \to 0 \tag{27}$$

Proof.

- **Learning Effect**: The forecast is updated whenever the Poisson shock arrives, and it is given by:

$$\hat{\mu}_t^i = e^{-\lambda}\hat{\mu}_{t-1}^i + (1-e^{-\lambda})\frac{\sigma_u^2}{\sigma_u^2+\gamma^2}\mu_t^i$$

For very large σ_u, we have that forecast errors evolve as

$$\begin{aligned}
f_t(z) = \mu_t(z) - \hat{\mu}_t(z) &= \mu_t(z) - e^{-\lambda}\hat{\mu}_{t-1}(z) - (1-e^{-\lambda})\frac{1}{1+\frac{\sigma_u^2}{\gamma^2}}\mu_t(z) \\
&= e^{-\lambda}\left(\mu_t(z) - \hat{\mu}_{t-1}(z)\right) \\
&= e^{-\lambda}\left(f_{t-1}(z) + \mu_t(z) - \mu_{t-1}(z)\right)
\end{aligned}$$

Taking the cross-sectional average:

$$\mathcal{F}_t = \int_0^1 f_t(z)dz = e^{-\lambda}\mathcal{F}_{t-1} + e^{-\lambda}\int_0^1 \underbrace{\left[\mu_t(z) - \mu_{t-1}(z)\right] dz}_{=0} = e^{-\lambda t}\delta$$

The half-life of the forecast errors is $\ln 2 / \ln\left(\frac{1+(\sigma_u/\gamma)^2}{1+(1-\lambda)(\sigma_u/\gamma)^2}\right)$ which is increasing in the ratio of noise to fundamental volatility γ/σ_u and decreasing in λ.

- **Selection Effect**: Now, to compute the mean markup estimation we condition on receiving a Poisson shock or not:

$$\begin{aligned}
\mathbb{E}_i[\hat{\mu}_t^i] &= (1-e^{-\lambda})\mathbb{E}_i[\hat{\mu}_t^i | J_t = 1] + e^{-\lambda}\mathbb{E}_i[\hat{\mu}_t^i | J_t = 0] \\
&\approx (1-e^{-\lambda})\mathbb{E}_i\left[\left(\frac{\sigma_u^2}{\sigma_u^2+\gamma^2}\right) F_{t-1}^i - \Delta p_{t-1}^{\delta,i} \Big| J_t = 1\right] + e^{-\lambda}\mathbb{E}_i[\hat{\mu}_{t-1}^i] \\
&\approx 0 + e^{-\lambda}\mathbb{E}_i[\hat{\mu}_{t-1}^i] \qquad (28)
\end{aligned}$$

In the second line, we use that (i) firms that do not receive a shock do not adjust their prices and do not update their forecasts, and (ii) firms that receive a Poisson shock with large variance (σ_u is large) will update their forecast and also will change their price in the same amount of the forecast update, canceling each other. Finally, since the initial condition for $\mathbb{E}_i[\hat{\mu}_{t-1}^i]$ is close to zero, $\mathbb{E}_i[\hat{\mu}_t^i]$ is also close to zero.

■