# Firm Uncertainty Cycles and the Propagation of Nominal Shocks

*Online Appendix: Not For Publication*

**Isaac Baley*** **Julio A. Blanco** †

March 23, 2018

## Contents



*Universitat Pompeu Fabra and Barcelona GSE, isaac.baley@upf.edu
†University of Michigan, jablanco@umich.edu

# A Markup shocks

## A.1 Leptokurtic Risk at the Firm-Level

In order to match micro-price statistics, the parametrization of our model uses a large volatility of the infrequent shocks ($\sigma_u^2$). Such calibration implies that markups' growth rate has a leptokurtic distribution. A leptokurtic distribution has heavy tails, which means that more of the variance is the result of infrequent extreme deviations as opposed to frequent modestly sized deviations, as in a Gaussian distribution. In this section, we provide empirical evidence of leptokurtic growth rates in several firm variables: profit rate, employment, sales, and capital. We think of this evidence as suggestive of the presence of large infrequent shocks affecting firms, which in turn generate leptokurtic distributions of firm outcomes.

### A.1.1 Variables' Definition

We use annual data from COMPUSTAT for publicly traded firms for the period between 1980 to 2015. Here we describe how we construct the variables of interest and their corresponding acronyms in COMPUSTAT.

- **Profits rate changes**: We define the profit rate as Operating Income Before Depreciation (*oibdp*) divided by a measure of capital given by Property, Plant and Equipment (*ppent*): $\pi_{ti} \equiv oibdp_{ti}/ppent_{ti}$. The profit rate changes are given by

$$\Delta\pi_{ti} = \pi_{ti} - \pi_{t-1i} \tag{A.1}$$

- **Employment growth rate**: We used Number of People Employed (*emp*) as a measure of employment ($n_{ti} \equiv emp_{ti}$), and define employment growth rate as

$$\Delta n_{ti} = \frac{emp_{ti} - emp_{t-1i}}{0.5 * (emp_{ti} + emp_{t-1i})} \tag{A.2}$$

- **Sales growth rate**: We use the Sales/Turnover ratio (*sale*) as a measure of sales ($y_{ti} = sale_{ti}$), and define sales growth rate as

$$\Delta y_{ti} = \frac{sale_{ti} - sale_{t-1i}}{0.5 * (sale_{ti} + sale_{t-1i})} \tag{A.3}$$

- **Capital growth rate**: We use Property, Plant and Equipment (*ppent*) as a measure of capital ($k_{ti} = ppent_{ti}$), and Capital Expenditures over Sale of Property ($capx - sppe$) as a measure of investment. We define capital growth as

$$\Delta k_{ti} = \frac{(capx - sppe)_{ti}}{0.5 * (ppent_{ti} + ppent_{t-1i})} \tag{A.4}$$

**Normalized growth rate** For each $x \in \{\pi, n, y, k\}$, we construct a normalized growth rate as follows:

$$\Delta\tilde{x}_{ti} = \Delta x_{ti} - \sum_i \frac{\Delta x_{ti}}{I(t)} - \sum_t \frac{\Delta x_{ti}}{T(i)} \tag{A.5}$$

where $I(t)$ is the average growth rate across firms in period $t$, and $T(i)$ is the average growth rate across periods for firm $i$. This normalized growth rate allows us to control for aggregate fluctuations and heterogeneity across firms.

**Data cleaning** We drop observations with negative total assets, negative capital or negative sales; firms with less than 5 years; firms operating in financial or regulated sectors (from SIC); and the $\pm 1\%$ outliers in growth rates.

**Weights** For each variable, we compute unweighted and weighted statistics [w], where we consider Total Assets (*atq*) as weights in order to control for size.

Table I: COMPUSTAT Raw Data Description

| **Label** | **Short description** | **Unit of measure** |
|---|---|---|
| gvkey | Firm Identifier | - |
| sic | Standard Industry Classification | - |
| oibdp | Operating Income Before Depreciation | Millions of dollars |
| ppent | Property, Plant and Equipment - Total (Net) | Millions of dollars |
| emp | Employees | Thousand |
| capx/sppe | Capital Expenditure/Sale of Property | Millions of dollars |
| sale | Sales/Turnover (Net) | Millions of dollars |
| atq | Asset Total | Millions of dollars |

All variables measured at annual frequency.
Link: https://wrds-web.wharton.upenn.edu/wrds/ds/crsp/ccm_a/funda/index.cfm?navId=120;

### A.1.2 Cross-Sectional Moments

Table II computes cross-sectional moments of the normalized growth rates. We find that after controlling for aggregate fluctuations, heterogeneity in the growths across firms, and firms' relative size, all variables' growth rates are largely leptokurtic, with kurtosis above 6 for all unweighted moments and above 10 for all weighted moments (recall that the benchmark level of kurtosis is 3 for a Normal random variable).

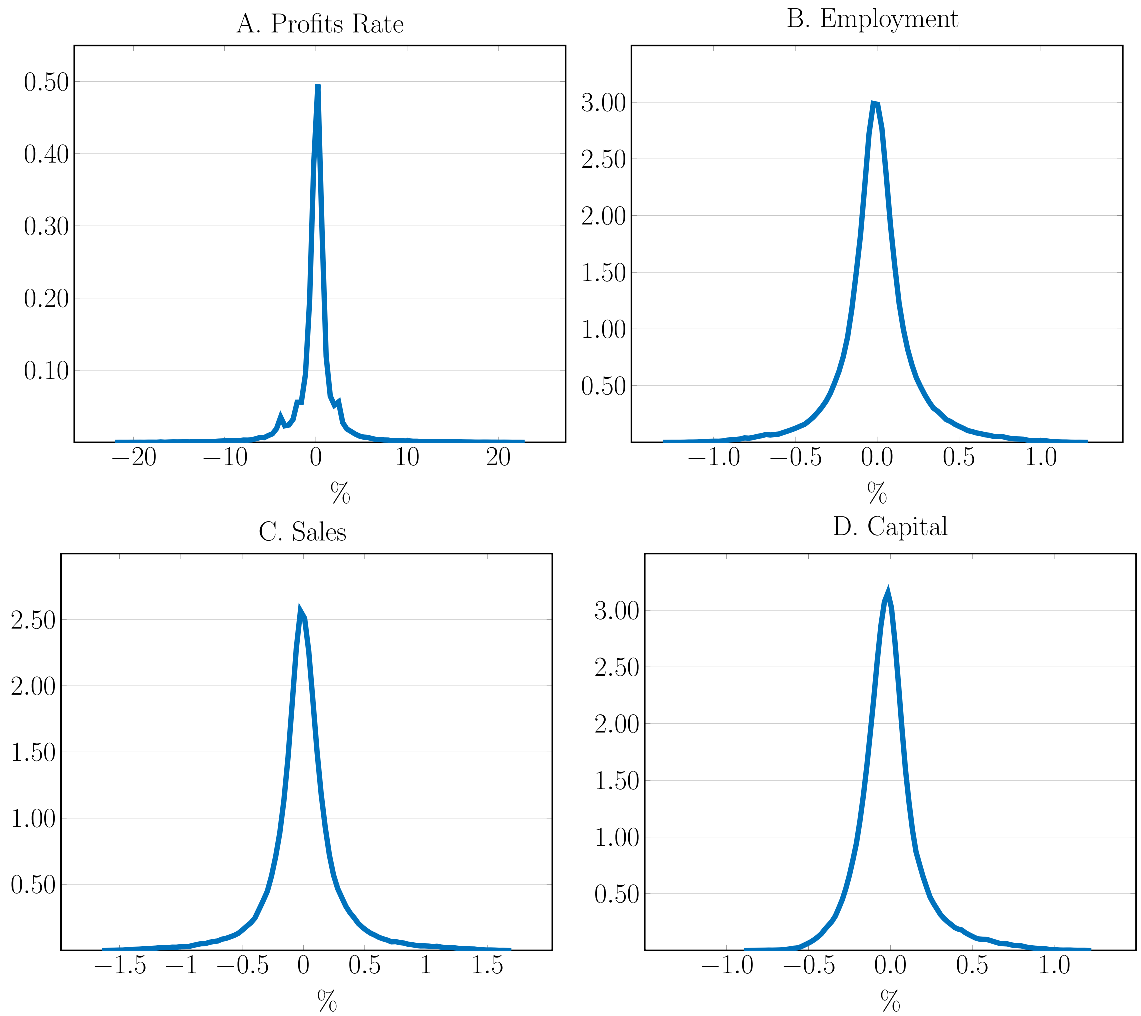
Table II: Cross-Sectional Moments of Variables' Growth Rates

| **Moment** | **Profits** | **Employment** | **Output** | **Capital** |
|---|---|---|---|---|
| **Unweighted Cross-Sectional Moments** | | | | |
| Mean | 0.039 | -0.000 | 0.001 | -0.001 |
| Median | 0.098 | -0.005 | -0.004 | -0.018 |
| Standard Deviation | 2.578 | 0.231 | 0.294 | 0.197 |
| Skewness | 0.295 | 0.122 | 0.134 | 1.007 |
| Kurtosis | 15.547 | 6.286 | 7.436 | 6.125 |
| **Weighted Cross-Sectional Moments** | | | | |
| Mean [w] | -0.035 | 0.002 | 0.004 | -0.002 |
| Median [w] | 0.078 | -0.002 | -0.004 | -0.005 |
| Standard Deviation [w] | 1.337 | 0.150 | 0.179 | 0.101 |
| Skewness [w] | -0.170 | 0.483 | 0.539 | 0.699 |
| Kurtosis [w] | 30.324 | 11.174 | 11.097 | 10.488 |
| **Correlations** | | | | |
| Corr. with sales | 0.005 | 0.442 | 1.000 | 0.232 |
| Corr. with sales [w] | 0.056 | 0.478 | 1.000 | 0.223 |
| **Persistence** | | | | |
| AR(1) coefficient | -0.036 | 0.050 | 0.069 | 0.193 |
| AR(1) coefficient [w] | 0.033 | 0.060 | 0.149 | 0.455 |

### A.1.3 Kernel Densities

Figure I plots the kernel estimate of the growth rate density for each variable. These figures provide a visual corroboration of the cross-sectional moments computed above. Densities are slender and fat-tailed, i.e. leptokurtic.

Figure I: Kernel Density of Profit Rate, Employment, Sales and Capital Growth



## A.2 Alternative interpretation for markup-gap shocks

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

This section derives the stochastic process of the markup-gap process and shows that it can be driven either by demand shocks—shocks to the demand elasticity with homothetic preferences—or by supply shocks—technology shocks to firm-level productivity. Following the environment proposed by Woodford (2009) and used by Álvarez, Lippi and Paciello (2011), Kehoe and Midrigan (2015), Midrigan (2011) and others, we specify preferences and technology to derive the following demand function $y_{it}^d(p_{it})$, cost function $c_{it}(y_{it})$, and the real profit function $\pi_{it}(p_{it})$:

$$y_{it}^d(p_{it}) = C_t A_{it} \left(\frac{p_{it}}{P_t}\right)^{-\gamma_{it}} \tag{A.6}$$

$$c_{it}(y_{it}) = \frac{y_{it} E_t}{A_{it}} \tag{A.7}$$

$$\pi_{it}(p_{it}) = P_t^{-1} y_{it}^d(p_{it})(p_{it} - c_{it}(y_{it}^d(p_{it})) \tag{A.8}$$

where $C_t$ is aggregate consumption, $A_{it}$ is a quality shock, $p_{it}$ is the firm price, $P_t$ is the aggregate price level, $\gamma_{it}$ is the demand elasticity of good $i$, and $E_t$ is the nominal wage. We can write the profits in term of markups and we have that

$$\pi_{it}(\mu_{it}) = C_t \left(\frac{E_t}{P_t}\right)^{1-\gamma_{it}} \mu_{it}^{\gamma_{it}} (\mu_{it} - 1) \tag{A.9}$$

where the markup is given by $\mu_{it} = \frac{p_{it} E_t}{A_{it}}$. The markup-gap is defined as $\log(\mu_{it}/\mu_{it}^*)$ where $\mu_{it}^*$ is the optimal static markup given by $\mu_{it}^* = \frac{\gamma_{it}}{\gamma_{it}-1}$. In a steady state with constant nominal wages $E$, whenever the firm does not change the price, the stochastic process of the markup-gap is given by

$$d\log(\mu_{it}/\mu_{it}^*) = -d\log(A_{it}) + d\log\left(\frac{\gamma_{it}}{\gamma_{it}-1}\right) \tag{A.10}$$

Therefore, it is equivalent to assume a process for log productivity or for the demand elasticity, as both translate into a process for the markup gap, i.e. either $d\log(A_{it}) = \sigma_f dW_t + \sigma_u u_t dQ_t$ or $d\log\left(\frac{\gamma_{it}}{\gamma_{it}-1}\right) = \sigma_f dW_t + \sigma_u u_t dQ_t$ delivers the stochastic process in the main text.

## B Comments on Filtering

### B.1 Alternative Filters for Non-Gaussian Models

The key assumptions to have a finite state representation in bayesian updating are Normality and linearity, see Davis (1977). The compound Poisson process in our model breaks the former assumptions, thus the firm's problem has an infinite state space. This is impossible to solve even numerically. The crucial assumption that we made to address this challenge is that the firm knows the arrival of a regime change. This allows us to keep the problem within a finite dimensional state Gaussian framework, as we show in Proposition 1, where only the first two moments of posterior distributions are needed for the firm's decision problem. Other approaches proposed in the literature that are infeasible in our framework are the following:

- *Discretization of the state space*: This approach was develop by Hamilton (1989). The main idea is to discretize the firm's state space and keep track the probability of being in each bin. This is infeasible, even numerically, since to capture a realistic stochastic process for markup and therefore the option value effect there has to be at least 5 points (zero, two symmetric points in the interior of Ss bands and two symmetric points in the exterior of the Ss bands). This will imply a five dimensional state space and the course of dimensionality.

- *Particle filter*: This method to estimate non-linear model is unfeasible due to the infinite space state and the computational intensity.

- *Approximation of the Posterior*: The first approximation is to ignore jump-process in the filtering equations and approximate the compound poisson process with a Brownian motion with variance given by $\lambda\sigma_u^2$ and apply the standard Kalman filter. This approach undoes the mechanism of state dependency in the estimation and the state dependency of our filtering. There are different types of approximation. For example, one approximation is proposed in Kim (1994). The main idea is to estimate an approximation of the current regime. This method has a very large state space or a short memory to update the state.

### B.2 Technical Discussion over the Filter

This section discusses the stochastic process of the markup-up gap and the markup-gap estimate. The state stochastic process is given by

$$d\mu_t = \sigma_f dW_t + \sigma_u u_t dQ_t \tag{B.11}$$

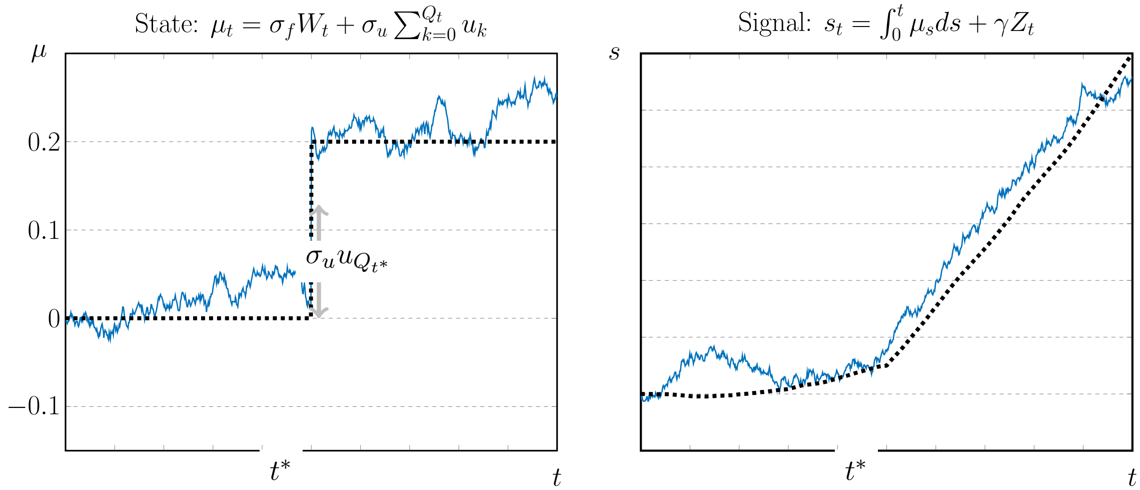
$$ds_t = \mu_t dt + \gamma dZ_t \tag{B.12}$$

where $W_t$ is a Wiener process, $u_tQ_t$ is a compound Poisson process with the Poisson counter's intensity $\lambda$, and $\sigma_f$ and $\sigma_u$ are the respective volatilities. When $dQ_t = 1$, the markup gap receives a Gaussian innovation $u_t \sim \mathcal{N}(0,1)$. The process $Q_t$ is independent of $W_t$ and $u_t$.

Figure II illustrates the evolution of the markup gap and the signal process. It assumes that there is a regime change at time $t^*$. At that moment, the average *level* of the markup gap jumps to a new value; nevertheless, the signal has continuous paths and only its *slope* changes to a new average value. This continuity together with the zero mean of $u_t$ implies continuous path for the markup-gap estimates.

It is also worth noticing that both the filtered estimates $\mu_t|\mathcal{I}_t$ and smoothed estimates $\mu_{t-\delta}|\mathcal{I}_t$ with $\delta > 0$ are Gaussian. In contrast, the predicted estimate $(\mu_{t+\delta}|\mathcal{I}_t)$ is not. For instance, in the case $\sigma_f = 0$, the predicted markup converges to a Laplace distribution with fat tails. We focus our

Figure II: Illustration of the Markup Gap and the Signal Processes




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

attention on the filtered estimate since it is the only input in our firm's decision problem. We leave for further research the analysis of other estimates.

## B.3 Filter and Policy Function with Non-Zero Mean Compound Process

This section characterizes the filter and the policy function whenever there $u_t$ has a non-zero mean $\bar{u}$.

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

$$\begin{aligned}
(state) \quad & d\mu_t &=& \; F\mu_t dt + \sigma_f dW_t + \sigma_u u_t dQ_t \\
(observation) \quad & ds_t &=& \; G\mu_t dt + \gamma dZ_t \\
(initial\ conditions\ for\ state) \quad & \mu_0 &\sim& \; \mathcal{N}(a, b) \\
(initial\ conditions\ for\ observations) \quad & s_0 &=& \; 0 \\
where \quad W_t, Z_t \sim Wiener\ Process, \quad & Q_t &\sim& \; Poisson(\lambda), \quad u_t \sim \mathcal{N}(\bar{u}, 1)
\end{aligned} \tag{B.13}$$

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

$$\begin{aligned}
d\hat{\mu}_t &= \left(F - \tfrac{G^2 \Sigma_t}{\gamma^2}\right) \hat{\mu}_t dt + \tfrac{G\Sigma_t}{\gamma^2} ds_t + \bar{u} dQ_t, \qquad \hat{\mu}_0 = a \\
d\Sigma_t &= \left(2F\Sigma_t + \sigma_f^2 - \tfrac{G^2 \Sigma_t^2}{\gamma^2}\right) dt + \sigma_u^2 dQ_t, \quad \Sigma_0 = b
\end{aligned} \tag{B.14}$$

To characterize the policy in our specific example, we use the values $F = 0$ and $G = 1$ we have the new representation of the state space

$$\begin{aligned} d\hat{\mu}_t &= \Omega_t d\hat{Z}_t + \bar{u} dQ_t \\ d\Omega_t &= \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t \end{aligned} \tag{B.15}$$

Next we characterize the policy function whenever $\bar{u}$ is sufficiently large or sufficiently small. The case where $\bar{u}$ is sufficiently small, the effect of $\bar{u}$ in the policy can be ignored up to a first order; and the case where $\bar{u}$ is sufficiently large, the policy is independent of $\bar{u}$ and similar to a mean reverting process for the markup-gap estimat.

**Proposition 2** *Let $\bar{\mu}^{\bar{u}}(\Omega)$ be the boundary of the continuation region whenever the markup-gap estimate and uncertainty follows B.15. Let $\max_\Omega \bar{\mu}^0(\Omega) \leq A$, then $\bar{\mu}^A(\Omega) = \bar{\mu}^{\bar{u}}(\Omega)$ for all $|\bar{u}| > A$ and can be approximated by*

$$V(\hat{\mu}, \Omega) = \max_\tau \mathbb{E}\left[\int_0^\tau e^{-rt} \hat{\mu}^2 dt + e^{-r\tau}\left(-\theta + V(0, \Omega_\tau)\right)\right] \tag{B.16}$$

$$d\hat{\mu}_t = -\lambda \hat{\mu} dt + \Omega_t d\hat{Z}_t \tag{B.17}$$

$$d\Omega_t = \frac{\sigma_f^2 + \lambda \sigma_u^2 - \Omega_t^2}{\gamma} dt \tag{B.18}$$

**Proof.** Assume For any $\bar{u}$ In the interior of the inaction region the value function solves $V$ solves the HJB equation:

$$rV^{\bar{u}}(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 - \Omega^2}{\gamma}\right) V^{\bar{u}}\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} V^{\bar{u}}_{\hat{\mu}^2}(\hat{\mu}, \Omega) \tag{B.19}$$

$$+ \lambda\left[\max\left\{V^{\bar{u}}\left(0, \Omega + \frac{\sigma_u^2}{\gamma}\right) - \theta, V^{\bar{u}}\left(\hat{\mu} + \bar{u}, \Omega + \frac{\sigma_u^2}{\gamma}\right)\right\} - V^{\bar{u}}(\hat{\mu}, \Omega)\right] \tag{B.20}$$

the value matching and the smooth pasting condition given by

$$V^{\bar{u}}(0, \Omega) - \bar{\theta} = V^{\bar{u}}(\pm\bar{\mu}(\Omega), \Omega) \tag{B.21}$$

$$V^{\bar{u}}_{\hat{\mu}}(\pm\bar{\mu}(\Omega), \Omega) = 0, \quad V^{\bar{u}}_\Omega(\pm\bar{\mu}(\Omega), \Omega) = V^{\bar{u}}_\Omega(0, \Omega) \tag{B.22}$$

Using the assumption that $|\bar{u}| > A$ we have that

$$rV^{\bar{u}}(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 - \Omega^2}{\gamma}\right) V^{\bar{u}}\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} V^{\bar{u}}_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda\left[V^{\bar{u}}\left(0, \Omega + \frac{\sigma_u^2}{\gamma}\right) - V^{\bar{u}}(\hat{\mu}, \Omega)\right] \tag{B.23}$$

Therefore

$$(\lambda + r)V^{\bar{u}}(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 - \Omega^2}{\gamma}\right) V^{\bar{u}}_\Omega(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} V^{\bar{u}}_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda V^{\bar{u}}\left(0, \Omega + \frac{\sigma_u^2}{\gamma}\right) \tag{B.24}$$

Notice that the policy is independent of $\bar{u}$. Moreover up to first order $V^{\bar{u}}\left(0, \Omega + \frac{\sigma_u^2}{\gamma}\right) = V^{\bar{u}}(\mu, \Omega) - V^{\bar{u}}_\mu(\mu, \Omega)\hat{\mu} + V^{\bar{u}}_\Omega(\mu, \Omega)\frac{\Omega}{\gamma}$, thus

$$rV(\hat{\mu}, \Omega) = -\hat{\mu}^2 + \left(\frac{\sigma_f^2 \lambda \sigma_u^2 - \Omega^2}{\gamma}\right) V_\Omega(\hat{\mu}, \Omega) - \lambda \hat{\mu} V_{\hat{\mu}}(\mu, \Omega) + \frac{\Omega^2}{2} V_{\hat{\mu}^2}(\hat{\mu}, \Omega) \tag{B.25}$$

Thus we have the result. ■

The main intuition of the previous proposition is simple: large changes in the markup-gap trigger price changes and therefore a drift in the markup-gap estimate toward zero. See Álvarez and Lippi (2014) for more details in the characterization of the function with respect to this stochastic process.

**Proposition 3** *Let $\mu^{*,\bar{u}}(\Omega)$ be the reset markup and $\bar{\mu}^{\bar{u}}(\Omega)$ the positive boundary of the continuation region. Then $\bar{\mu}^{\bar{u}}(\Omega) - \mu^{*,\bar{u}}(\Omega) = \bar{\mu}(\Omega) + o(\Omega^2)$ where $\bar{\mu}(\Omega)+$ is the policy in the main text.*

**Proof.** Let $\bar{\mu}^{\bar{u}}(\Omega)$ the positive boundary of the continuation region and $\underline{\mu}^{\bar{u}}(\Omega)$ the negative boundary of the continuation region. Given the symmetry of the problem we have that $\bar{\mu}^{\bar{u}}(\Omega) - \mu^{*,\bar{u}}(\Omega) = -(\underline{\mu}^{-\bar{u}}(\Omega) - \mu^{*,-\bar{u}}(\Omega))$, thus $\bar{\mu}^{\bar{u}}(\Omega) - \mu^{*,\bar{u}}(\Omega)$ is symmetric over $\bar{u}$ around 0 and there is zero order effect over $Eu$. ■

The main intuition of the previous proposition is simple: positive or negative value of $\bar{u}$ have second order effect in the width of the Ss bands since they are similar to changes in the drift. This shock only changes the level of the domain of the markup-up gap estimate.

### B.4 Discrete time approximation (for simulation)

Consider a small time interval $\Delta$. Let $b_t^W$ and $b_t^Z$ be two binomial random variables that take values of 1 and -1 with probability 1/2 and $q_t^Q$ be an additional binomial random variable that takes value 1 with probability $\lambda\Delta$ and a value 0 otherwise. Using these elements, we can approximate the continuous laws of motion above as follow (note that it is convenient to track desired price changes $\Delta p$, which only materialize when they fall outside the inaction region):

$$\Delta p = \hat{\mu}_{t-\Delta} - \frac{\Omega_{t-\Delta}}{\gamma}\hat{\mu}_{t-\Delta}\Delta + \frac{\Omega_{t-\Delta}}{\gamma}\left(s_t - s_{t-\Delta}\right) \tag{B.26}$$

$$\mu_t = \mu_{t-\Delta} + \sigma_f\sqrt{\Delta}b_t^W + \sigma_u u_t q_t^Q - \Delta p \; \mathbb{1}_{\{|\Delta p| > \bar{\mu}(\Omega)\}}, \tag{B.27}$$

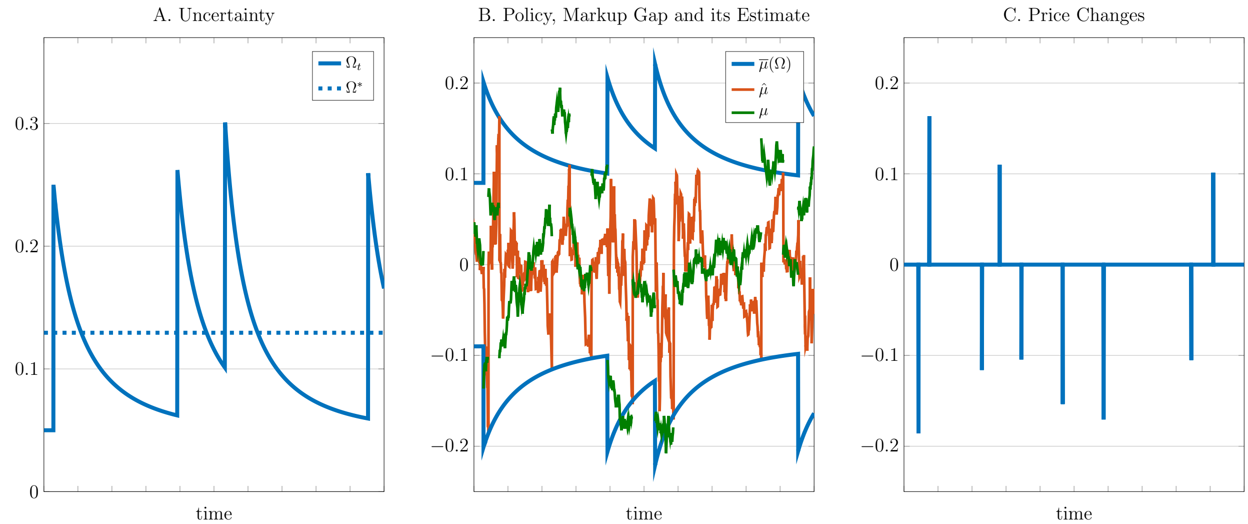
$$s_t = s_{t-\Delta} + \mu_{t-\Delta}\Delta + \gamma\sqrt{\Delta}b_t^Z, \tag{B.28}$$

$$\hat{\mu}_t = \Delta p \; \mathbb{1}_{\{|\Delta p| \leq \bar{\mu}(\Omega)\}}, \tag{B.29}$$

$$\Omega_t = \Omega_{t-\Delta} + \frac{\sigma_f^2 - \Omega_{t-\Delta}}{\gamma}\Delta + \frac{\sigma_u^2}{\gamma}q_t^Q \tag{B.30}$$

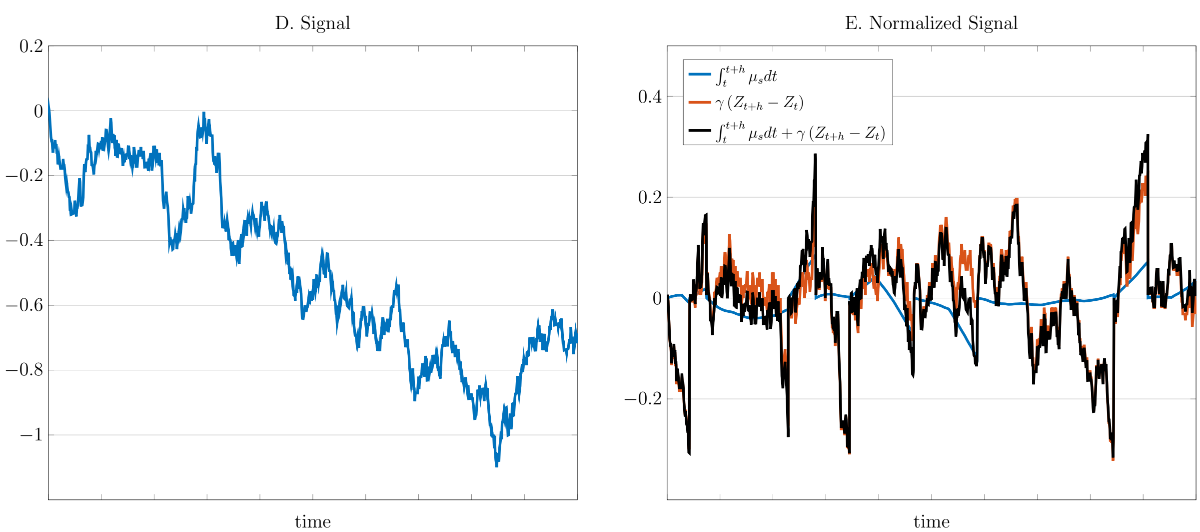
The left panel in the next figure shows the evolution of the signal process $s_t$; its slope tracks the true markup. In the right panel, we also plot the signal but reseting it to zero every time the price is adjusted, i.e. we show the the accumulated variation in the signal between price changes. In the same figure, we also show the two signal components, which are its trend (the true markup), and the noise process $\gamma dZ_t$.

Figure III: Sample Paths For One Firm




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

Figure IV: Sample Paths For Signals




Panel A: Signal. Panel B: Normalized Signal. This figure simulates one realization of the stochastic processes using the finite difference method and uses the analytical approximation of the inaction region.

# C Infinitesimal Generator and Adjoint Operator

This section derives the infinitesimal generator and its adjoint operator. These two functional equations are use to characterize the policy, micro-price statistics and the effect of a money shock.

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

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

For our problem, the generator is given by:

$$\mathcal{A}\phi(\hat{\mu}_t, \Omega_t) = \frac{\sigma_f^2 - \Omega_t^2}{\gamma}\phi_\Omega(\hat{\mu}_t, \Omega_t) + \frac{\Omega_t^2}{2}\phi_{\hat{\mu}^2}(\hat{\mu}_t, \Omega_t) + \lambda\left[\phi\left(\hat{\mu}_t, \Omega_t + \frac{\sigma_u^2}{\gamma}\right) - \phi(\hat{\mu}_t, \Omega_t)\right] \quad \text{(C.31)}$$

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

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

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

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

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

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

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

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

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

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

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

$$\mathcal{A}^* f(\hat{\mu}, \Omega) = -\frac{\sigma_f^2 - \Omega^2}{\gamma} f_\Omega(\hat{\mu}, \Omega) + \frac{2\Omega}{\gamma} f(\hat{\mu}, \Omega) + \frac{\Omega^2}{2} f_{\hat{\mu}^2}(\hat{\mu}, \Omega) + \lambda \left[ f\left(\hat{\mu}, \Omega - \frac{\sigma_u^2}{\gamma}\right) - f(\hat{\mu}, \Omega) \right] \tag{C.32}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

■

# D Conditions for the Optimality of the Continuation Region

The proof of our Proposition 3 applies Theorem 2.2 in Øksendal and Sulem (2010) to obtain sufficient conditions that characterize the value function and the optimal policy, i.e. the HJB equation, value matching and smooth pasting conditions. Our strategy for characterizing the policy function is to impose the SPC following Theorem 2.2 in Øksendal and Sulem (2010) (or alternatively, Theorem 10.4.1 in Øksendal 2007 for diffusion processes) assuming the necessary conditions for the theorem hold, and then we check the validity of such assumptions in the Online Appendix Section E.3. In this section we provide further detail regarding the conditions needed for that Theorem to apply. We also provide some examples that illustrate the conditions. Importantly, the theorem does not require Brownian Motion in any dimension, but imposes other conditions. Among these, the key condition related to the SPC that needs to be verified is that the problem has a **zero local measure at the boundary of the continuation region (ZLM)**. This condition imposes **regularity** of the stochastic process at the boundary of the continuation set. Lastly, we check numerically the validity of the value matching and smooth pasting conditions by solving exactly the value function using a finite difference method (see Section E.3).

## D.1 Setup

In order to apply Theorem 2.2 in Øksendal and Sulem (2010), we need to extend the stochastic process of the state to include time. Let $Y_t = (\hat{\mu}_t, \Omega_t, z_t)$ be a jump diffusion process given by

$$
\begin{aligned}
d\hat{\mu}_t &= \Omega_t d\hat{Z}_t && \text{(D.33)}\\
d\Omega_t &= \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t && \text{(D.34)}\\
dz_t &= dt && \text{(D.35)}
\end{aligned}
$$

where the third element is just capturing the evolution of time. Let $\mathcal{F}_t$ be the filtration generated by the previous stochastic process and let $\mathcal{T}$ be the set of stopping times with respect to $\mathcal{F}_t$. Finally, let $f$ and $g$ be the payoff and terminal functions respectively

$$
\begin{aligned}
f(Y_t) &= -e^{-rz_t}\hat{\mu}_t^2 dt && \text{(D.36)}\\
g(Y_t) &= e^{-rz_t}\left(-\bar{\theta} + V(0, \Omega_t)\right) && \text{(D.37)}
\end{aligned}
$$

Then the firm problem is given by

$$
V(\hat{\mu}_0, \Omega_0) = \max_{\tau \in \mathcal{T}} \mathbb{E}\left[\int_0^{\tau} -e^{-rz_s}\hat{\mu}_s^2 ds + e^{-rz_\tau}\Big(-\bar{\theta} + V(0, \Omega_\tau)\Big)\Big|\mathcal{I}_0\right] \quad \text{(D.38)}
$$

subject to the stochastic process above. Note that we have replaced $z_t$ in the payoff and terminal functions. For simplicity we have already used that the optimal control (reset markup gap) is $\mu' = 0$.

## D.2 Theorem

Now we enunciate Theorem 2.2 in Øksendal and Sulem (2010).

**Theorem 1** *Let $S$ be a set. Suppose that we can find a function $\phi : \bar{S} \to \mathbb{R}$ such that*

1. $\phi \in C^1(S) \cap C(S)$ .
2. $\phi \geq g$.

*Define the continuation region* $D \equiv \{y \in S : \phi(y) > g(y)\}$ *and define the first exit time* $\tau_D \equiv \inf\{t > 0 : Y_t \notin D\} < \infty \quad a.s. \ \forall y$. *Suppose the following conditions also hold:*

3. $\mathbb{E}^y \left[\int_0^{\tau_S} \mathbb{1}_{\partial D}(Y_t) dt\right] = 0$ *(zero local mass at the boundary of the continuation region).*

4. $\mathcal{A}\phi + f = 0$ *on* $D$*, where* $\mathcal{A}$ *is the infinitesimal generator.*

5. *The boundary of the continuation region* ($\partial D$) *is a Lipschitz surface.*

6. $Y(\tau_S) \in \partial S$ *a.s. on* $\{\tau_S < \infty\}$ *and* $\lim_{t \to \tau_S^-} \phi(Y_t) = g(Y_{\tau_s}) \mathbb{1}_{\{\tau_S < \infty\}}$.

7. $\mathbb{E}^y \left[|\phi(Y_\tau)| + | \int_0^{\tau_S} |\mathcal{A}\phi(Y_t)| dt\right] < \infty$ *for all* $\tau \in \mathcal{T}$.

8. $\{\phi(Y_t); \tau \in \mathcal{T}, \tau \leq \tau_D\}$ *is uniformly integrable of all* $y$.

*Then* $\phi(y) = V(y)$*, i.e. the function* $\phi$ *solves the firm problem, and* $\tau_D$ *is an optimal stopping time.*

## D.3 Explanation of conditions

First of all, notice that the theorem states *sufficient conditions* to find the optimal policy, i.e. if a candidate function satisfies these conditions then we have found the optimal policy. This is what we state in the paper. Now, let us discuss each condition.

Condition 1 Establishes the smooth pasting condition by imposing differentiability of $\phi$ in the boundary $\partial D$. To see this, note that $\phi$ is differentiable outside $D$ since $\phi(y) = g(y) \forall y \notin D$ and $g(y)$ is differentiable. Also, the value function is differentiable in the interior $D^o$ by Condition 4. Therefore, by requiring differentiability of $\phi$ in $S$, it is requiring differentiability at the boundary of $D$.

Condition 2 It is related to the value matching condition.

Condition 3 The most important condition to check. We call it the **zero local measure at the boundary of the continuation region (ZLM)**. This condition imposes **regularity** of the stochastic process at the boundary of the continuation set. In words, the ZLM requires that the expected time the process spends in the boundary is zero, regardless of the point where the process started. Another way to think about this condition is that it requires a zero probability of positive expected duration at the boundary for any initial condition. Following the books' notation, this condition is stated as follows:

$$\mathbb{E}\left[\int_0^{\tau_G} \mathbb{1}_{\{\partial D\}}(Y_t)\ dt \middle| Y_0 = y\right] = 0 \quad \text{for all} \quad y \in G \tag{D.39}$$

where $y$ is the initial condition, $G$ is a domain that includes the continuation set, $\tau_G = \inf\{t \geq 0 : Y_t \notin G\}$ is a stopping time, and $\mathbb{1}_{\{\partial D\}}(Y_t)$ is an indicator function equal to one if the process $Y_t$ is at the boundary $\partial D$.

Condition 4 Requires differentiability of $\phi$ in D.

Condition 5 It says that there exists a function such that its graph is a Lipschitz surface.

Condition 6 Trivial if we assume $S$ is a closed set that contains the continuation region.

Condition 7 Discards infinite values of the value function.

Condition 8 It is a technical condition.

### D.4 Necessity of Condition 3 (ZLM)

Let us first illustrate the necessity of the ZLM through the means of two examples without Brownian motion. In the first example, we consider a deterministic problem with positive drift that satisfies the ZLM and thus the SPC holds too; this shows that Brownian motion is not needed. In the second example, we consider a jump process that does not satisfy the ZLM and thus the SPC does not hold; this shows the necessity of ZLM. Then we revisit our numerical checks. We hope that this analysis is clear and convinces you of our strategy.

**Example 1: ZLM and SPC hold without Brownian Motion**

- **Environment:** Consider a profit maximizing firm that chooses the price at which to sell its product, subject to inflation $\pi > 0$. To change the price, it must pay a menu cost $\theta$ in units of product. After the first price change, the firm disappears. Following the exposition in our main text, the problem can be described in terms of markup gaps $\mu_t$ as follows:

$$\begin{aligned} V(\mu) &= \max_{T\geq 0} \left\{ -\int_0^T e^{-\rho t}\mu_t^2 dt - e^{-\rho T}\theta \right\} \\ \mu_t &= \mu - \pi t, \qquad \pi > 0 \end{aligned} \tag{D.40}$$

- **ZLM:** Assume that the continuation region is a connected set $\mathcal{D} = [a, b]$. To verify that ZLM in (D.39) holds, notice that for any stopping time $T$ and initial condition $y$, the presence of the drift brings the stochastic process outside the boundary and thus its local measure is equal to zero. Formally, the following condition holds

$$\int_0^T \mathbb{1}_{\{\partial\mathcal{D}\}}(y - \pi t)dt = 0, \qquad \forall y, T \tag{D.41}$$

  Together, the positive drift and the countable number of point in the boundary of the continuation region imply a zero expected time at the boundary. It is important to remark that the ZLM holds only for $\pi > 0$. It is easy to verify that the ZLM does not hold for the case $\pi = 0$, since then the boundary would not be regular.

- **Characterization** Given that ZLM (and other assumptions) hold, we can apply Theorem 2.2 and obtain the following conditions for an optimal solution:
  - (i) a symmetric continuation region $\mathcal{D} = [-\sqrt{\rho\theta}, \sqrt{\rho\theta}]$, with $\partial\mathcal{D} = \pm\sqrt{\rho\theta}$,
  - (ii) the HBJ within the continuation region $\mathcal{D}$, namely $V'(\mu_t) = \frac{\mu_t^2 - \rho V(\mu_t)}{\pi}$ for all $\mu_t \in \mathcal{D}$,
  - (iii) the value matching condition: $V(\pm\sqrt{\rho\theta}) = -\theta$, and
  - (iv) the smooth pasting condition: $V'(\pm\sqrt{\rho\theta}) = 0$.

- **Verification of SPC from first principles**

  First, we obtain the optimal stopping time and the value function. Taking FOC in (D.40) we have that

$$\mu_{T(\mu)}^2 \geq \rho\theta \qquad \text{with equality if} \quad T(\mu) > 0. \tag{D.42}$$

  From the previous equation we obtain the optimal stopping time as a function of $\mu$

$$T(\mu) = \begin{cases} 0 & \text{if } |\mu| > \sqrt{\rho\theta} \\ \frac{\mu - \sqrt{\rho\theta}}{\pi} & \text{otherwise} \end{cases}. \tag{D.43}$$

Using the optimal stopping time, the value function is

$$V(\mu) = \begin{cases} -\theta & \text{if } |\mu| > \sqrt{\rho\theta} \\ -\left( \int_0^{\frac{\mu-\sqrt{\rho\theta}}{\pi}} e^{-\rho t} (\mu - \pi t)^2 \, dt + e^{-\rho\left(\frac{\mu-\sqrt{\rho\theta}}{\pi}\right)} \theta \right) & \text{otherwise} \end{cases} . \tag{D.44}$$

and its first derivative given by

$$V'(\mu) = \begin{cases} 0 & \text{if } |\mu| > \sqrt{\rho\theta} \\ -\int_0^{\frac{\mu-\sqrt{\rho\theta}}{\pi}} 2e^{-\rho t} (\mu - \pi t) \, dt + \frac{1}{\pi} e^{-\rho\left(\frac{\mu-\sqrt{\rho\theta}}{\pi}\right)} \underbrace{(\rho\theta - \rho\theta)}_{=0} & \text{otherwise} \end{cases} . \tag{D.45}$$

where the last term is always equal to zero by the envelope condition. It is easy to see that $V'(\mu)$ is continuous in $|\mu| > \sqrt{\rho\theta}$ and in $|\mu| < \sqrt{\rho\theta}$ (the Riemann Integral is continuous in the limit of integration).

Next, we show that the SPC holds. Since $V'(\mu)$ is constant outside the continuation region, it is continuous in such domain. To check the continuity at the boundary of the continuation region, we need to verify that the right and left limits at the boundary are equal. When approaching the left and right boundaries from outside, both the continuity is immediately verified:

$$\lim_{\mu \downarrow \sqrt{\rho\theta}} V'(\mu) = \lim_{\mu \uparrow -\sqrt{\rho\theta}} V'(\mu) = 0$$

And when approaching the left boundary from inside, we obtain

$$\lim_{\mu \downarrow -\sqrt{\rho\theta}} \int_0^{\frac{\mu-\sqrt{\rho\theta}}{\pi}} e^{-\rho t} (\mu - \pi t) \, dt = \lim_{\mu \downarrow -\sqrt{\rho\theta}} \left[ \frac{1 - e^{-\rho \frac{\mu-\sqrt{\rho\theta}}{\pi}}}{\rho} + \frac{1 - e^{-\rho \frac{\mu-\sqrt{\rho\theta}}{\pi}} (\rho \frac{\mu-\sqrt{\rho\theta}}{\pi} + 1)}{\rho^2} \right] = 0.$$

Analogously, it is shown that the limit when approaching the right boundary from inside is also zero, i.e. $\lim_{\mu \uparrow \sqrt{\rho\theta}} V'(\mu) = 0$. Since the limits from inside and outside the continuation region towards the boundary are equal, the SPC holds.

- **Discussion:** The SPC holds in this example even if we do not have a Brownian Motion for the state; moreover, the SPC is a necessary condition.

**Example 2: ZLM and SPC do not hold**

- **Environment:** Now idiosyncratic shocks follow a compound Poisson process $\phi_t dN_t$, where $dN_t$ is a counter with arrival rate $\lambda$ and $\phi_t$ are uniform innovations. For every price change, the firm must pay a menu cost $\theta$ in units of product. Again, following the exposition in the main text, the problem can be described in terms of markup gaps as follows:

$$\begin{aligned} V(\mu) &= \max_{T \geq 0} \mathbb{E} \left[ -\int_0^T e^{-\rho t} \mu_t^2 dt + e^{-\rho T} \left[ -\theta + \max_{\mu^*} V(\mu^*) \right] \right] \\ \mu_t &= \phi_t dN_t, \qquad \mu_0 = \mu \end{aligned}$$

- **ZLM:** To show that there is a positive local measure at the boundary, i.e. ZLM in (D.39) does not hold, let $\mathcal{D} = [a, b]$ be the continuation region and assume that the initial condition is exactly at the boundary, i.e. $y = a$. Then the expected time that the state will be at the boundary has a positive measure, because it will only get out if there is a Poisson shock. Formally, the expectation is bounded below by the probability that there is a Poisson shock:

$$\mathbb{E} \left[ \int_0^T \mathbb{1}_{\{\mathcal{D}\}}(\mu_t) \, dt \middle| \mu = a \right] \geq \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda} > 0$$

- **Optimal stopping time and value function** Since ZLM does not hold, we cannot apply Theorem 2.2 Øksendal and Sulem (2010). We derive the conditions by first principles. The HBJ is given by

$$(r+\lambda)V(\mu) = -\mu^2 + \lambda\mathbb{E}_\phi\left[\max\{V(0)-\theta, V(\mu+\phi)\}\right] \quad \text{for all} \quad \mu \in \mathcal{D} \tag{D.46}$$

Since $\phi$ is Uniform, the expectation $\Gamma \equiv \lambda\mathbb{E}_\phi\left[\max\{V(0)-\theta, V(\mu+\phi)\}\right]$ is independent of $\mu$. With this notation, the value function becomes

$$V(\mu) = \max\left\{\frac{-\mu^2+\Gamma}{r+\lambda}, -\theta + \frac{\Gamma}{r+\lambda}\right\}$$

with the continuation region given by $\mathcal{D} = \left[-\sqrt{\theta(r+\lambda)}, \sqrt{\theta(r+\lambda)}\right]$.

- **SPC:** The smooth pasting condition does not hold in this example, since the limits from outside and inside the continuation region towards its lower boundary are not equal:

$$\lim_{\mu\uparrow-\sqrt{\theta(r+\lambda)}} V'(\mu) = 0 \qquad \text{and} \qquad \lim_{\mu\downarrow-\sqrt{\theta(r+\lambda)}} V'(\mu) = 2\sqrt{\theta(r+\lambda)}. \tag{D.47}$$

Analogously, one can show that the limits do not coincide for the upper boundary.

- **Discussion:** In this example, the ZLM and SPC condition do not hold. Note, however, that if we added a drift in the process of $\mu$, then zero expected time at the boundary holds as in our previous example and Theorem 2.2 Øksendal and Sulem (2010) could be used.

## D.5 Check conditions for our problem

Now we check the conditions in the Theorem. To show these conditions we will use the solution to the value function in Proposition 4.

Condition 1 It is imposed for the solution, so it is satisfied.

Condition 2 Satisfied outside D by the value matching condition. To check in the interior of the continuation region, fix an $\Omega$. Moving in the $\hat{\mu}$ direction, the value function is decreasing in $|\hat{\mu}|$ until $\bar{\mu}(\Omega)$ where the value is $\phi(0,\Omega)-\theta$.

Condition 3 This is the problematic condition in the previous examples but in our model it holds. Notice that

$$\mathbb{E}\left[\int_0^{\tau_S} \chi_{\partial D}(Y_t)dt \Big| \hat{\mu}_0, \Omega_0\right] = \int_0^\infty Pr^{\hat{\mu}_0,\Omega_0}[Y_t = (\bar{\mu}_t(\Omega_t), \Omega_t)]dt$$

To show the previous condition, it is sufficient to show that the boundary is regular. Take any point s.t. $(\hat{\mu}_0, \Omega_0)$ satisfies $\hat{\mu}_0 = \bar{\mu}(\Omega_0)$. Then if $|\bar{\mu}'(\Omega_0)| < \infty$ the Brownian Motion for $\hat{\mu}$ generates $Pr^{\hat{\mu}_0,\Omega_0}[\tau_D = 0] = 1$. Moreover if $\bar{\mu}'(\Omega_0) > 0$, the drift in $\Omega$ generates $Pr^{\hat{\mu}_0,\Omega_0}[\tau = 0] = 1$. Thus the boundary set is regular.

Condition 4 Notice that the function $f(\Omega) = (\mu(\Omega), \Omega)$ is Lipschitz continuous as long $\bar{\mu}'(\Omega)$ is bounded. This property holds in our solution.

Condition 5 It is imposed for the solution.

Condition 6 Trivially satisfied.

Condition 7 Satisfied by Condition 8 and discounting.

Condition 8 We need to check that

$$\left|\int_0^t \Omega_s dZ_t\right| \leq \bar{\mu}(\Omega_t)$$

for some finite $t$ a.s. Notice that there is an uncertainty level $\Omega^*$ such that $\frac{\sigma_f^2+\sigma_u^2-\Omega^*}{\gamma} < 0$, which implies $d\Omega_t < 0$ with probability $1 - o(dt^2)$. Therefore, checking the previous condition is equivalent to checking the following

$$\left|\int_0^t \Omega_s dZ_t\right| \leq \sup_{\Omega\in[\sigma_f,\Omega^*]} \bar{\mu}(\Omega)$$

Since $\Omega_t \geq \sigma_f$ the previous condition is satisfied if

$$\left|\int_0^t \sigma_f dZ_t\right| \leq \sup_{\Omega\in[\sigma_f,\Omega^*]} \bar{\mu}(\Omega)$$

for some finite $t$ a.s. Checking the previous condition is standard.

### D.6 Numerical Verification of Smooth Pasting Condition

Besides the proofs above, we also use numerical methods to check ex-post the validity of the smooth pasting condition. We use a finite difference method to compute the value function. Please see Section C.3 of the Online Appendix for all the details.

# E Accuracy of Analytical Approximations

In this section we show that the analytical characterization of policies and price statistics in continuous time are indeed good approximations. For this task, we solve the model in continuous time numerically and plot the analytical and numerical counterparts for different sets of parameters.

## E.1 Numerical solution of the model in continuous time

We implement the continuous time model following the finite difference method to approximate the differential equations.

1. Construct grids for $\hat{\mu}$ and $\Omega$. Fix the number of grid points for uncertainty, $n_\Omega$, and define $h \equiv \frac{\sigma_u^2}{\gamma n_\Omega}$. The respective grids are given by

$$\begin{aligned} S_{\hat{\mu}} &= \{0, \pm h, 2 \pm h, \ldots, \pm \bar{\mu}\} \\ S_\Omega &= \{\sigma_f, \sigma_f + h, \ldots, \bar{\Omega}\} \end{aligned}$$

2. Find a discrete stochastic process that is locally consistency to the continuous stochastic process given by:

$$d\mu_t = \Omega_t d\hat{Z}_t \tag{E.48}$$

$$d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma} dt + \frac{\sigma_u^2}{\gamma} dQ_t \tag{E.49}$$

Define $\delta(\Omega, h) \equiv \Omega^2 + \frac{\Omega^2 - \sigma_f^2}{\gamma} h$ and $\Delta(\Omega, h) \equiv \frac{h^2}{\delta(\Omega,h)}$. For simplicity, we write $\delta$ and $\Delta$ without their dependence on $(\Omega, h)$. The discrete stochastic process for $(\hat{\mu}', \Omega')|(\hat{\mu}, \Omega)$ is defined by

$$(\hat{\mu}', \Omega')|(\hat{\mu}, \Omega) = \begin{cases} (\hat{\mu} - h, \Omega) & \text{with prob. } \frac{\Omega^2}{2\delta}(1 - \lambda\Delta) \\ (\hat{\mu} + h, \Omega) & \text{with prob. } \frac{\Omega^2}{2\delta(}(1 - \lambda\Delta) \\ (\hat{\mu}, \Omega - h) & \text{with prob. } \frac{(\Omega^2 - \sigma_f^2)h}{\gamma\delta}(1 - \lambda\Delta) \\ (\hat{\mu}, \Omega + n_\Omega h) & \text{with prob. } \lambda\Delta \end{cases} \tag{E.50}$$

It is straightforward to check the local consistency of this process. For instance, for $\hat{\mu}$ process, we have that:

$$\begin{aligned} \mathbb{E}[\Delta\mu] &= (h - h)\frac{\Omega}{2\delta}(1 - \lambda\Delta) = 0 \\ \mathbb{E}[(\Delta\mu)^2] &= 2h^2 \left[\frac{\Omega^2}{2\delta}(1 - \lambda\Delta)\right] = \Omega^2\Delta - \lambda\Omega^2\Delta^2 = \Omega^2\Delta + o(\Delta) \end{aligned}$$

3. Let $V^C$ be the value of changing the markup, $V^{NC}$ the value of not changing and $V$ the upper envelope of the previous values. The recursive problem of the firms is given by:

$$\begin{aligned} V(\hat{\mu}, \Omega) &= \max\left\{\mathbb{E}[V^C(\hat{\mu}', \Omega_s')|(\hat{\mu}, \Omega)],\ \mathbb{E}[V^{NC}(\hat{\mu}', \Omega_s')|(\hat{\mu}, \Omega)]\right\} \\ V^C(\hat{\mu}, \Omega) &= -\bar{\theta} + \max_x \left[-x^2\Delta(\Omega, h) + e^{-r\Delta(\Omega,h)}v(x, \Omega)\right] \\ V^{NC}(\hat{\mu}, \Omega) &= -\hat{\mu}^2\Delta(\Omega, h) + e^{-r\Delta}V(\hat{\mu}, \Omega) \end{aligned}$$

Notice the timing assumption in the recursive formulation: in the first equation we have the maximum of the expectation instead of the expectation of the maximum, this is, the firm has to make the pricing decision before the realization of the shock.

## E.2 Calibration in continuous time

We solve the continuous time model using the same calibration of the discrete time model at weekly frequency, except for three parameters. First, the discount rate is set very small at $r = 0.001$ to approximate the continuous time approximation. This tiny discount rate makes it difficult to solve numerically for the optimal policy as the Bellman equation has trouble converging. By increasing the volatility of the frequent shocks to $\sigma_f = 0.05$ and decreasing the menu cost to $\theta = 0.005$ we obtain convergence and maintain the main features of the calibration, such as the size and frequency of price adjustment.

## E.3 Numerical Verification of Smooth Pasting Condition

Recall that smooth pasting conditions imply that

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

Using the finite difference method, we can check numerically that these conditions are indeed satisfied. Given the locally consistent process in (E.50), as the number of grid points $n$ grows, the time period $\Delta$ shrinks, and the discrete approximation of the value function converges to the continuous time value function. In the limit, the smooth pasting conditions should also hold:

$$\lim_{n\to\infty} V_{\hat{\mu}}(\bar{\mu}(\Omega), \Omega) = \lim_{n\to\infty} V_{\Omega}(\bar{\mu}(\Omega), \Omega) - V_{\Omega}(0, \Omega) = 0$$

For each grid point $\hat{\mu}_i^n \in S_{\hat{\mu}}^n$, the numerical (left) partial derivative with respect to markup gap estimate $\hat{\mu}$ is computed as

$$V_{\hat{\mu}}(\hat{\mu}_i, \Omega) = \frac{V(\hat{\mu}_i, \Omega) - V(\hat{\mu}_{i-1}, \Omega)}{\hat{\mu}_i - \hat{\mu}_{i-1}}, \qquad \forall \Omega \quad \forall n$$

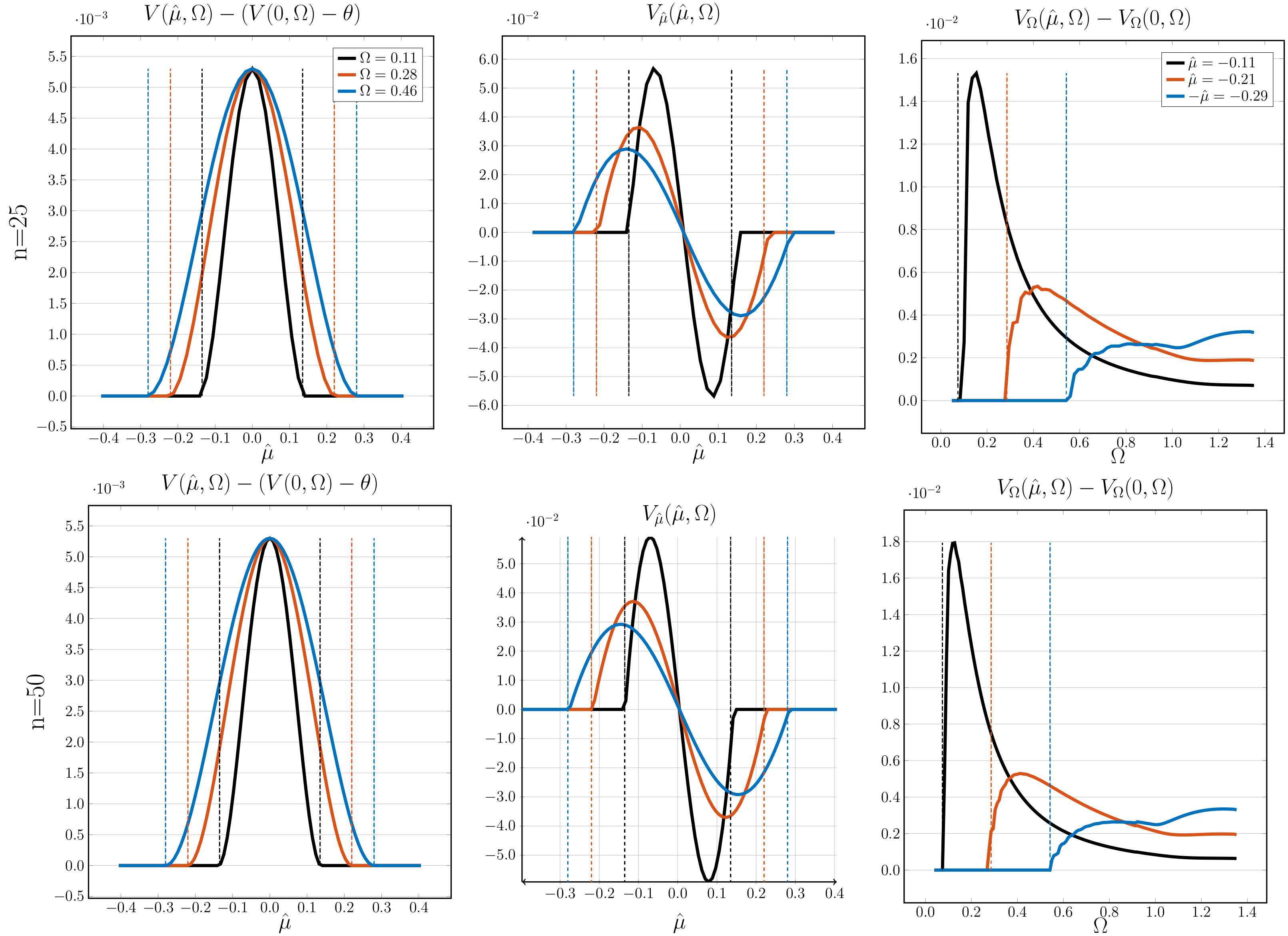
Analogously, for each grid point $\Omega_i^n \in S_{\Omega}$, the numerical (left) partial derivative with respect to uncertainty $\Omega$ is computed as

$$V_{\Omega}(\hat{\mu}, \Omega_i) = \frac{V(\hat{\mu}, \Omega_i) - V(\hat{\mu}, \Omega_{i-1})}{\Omega_i - \Omega_{i-1}} - \frac{V(0, \Omega_i^n) - V(0, \Omega_{i-1})}{\Omega_i - \Omega_{i-1}}, \qquad \forall \hat{\mu} \quad \forall n$$

Table III checks the smooth pasting condition at the border of the inaction region $\bar{\mu}$ for increasing values of $n$. The smooth pasting condition for $\hat{\mu}$ is evaluated at three values of uncertainty $\Omega \in \{0.11, 0.29, 0.46\}$, and the smooth pasting condition for $\Omega$ is evaluated at three values of the inaction border $-\bar{\mu} \in \{-0.3, -0.21, -0.11\}$. To avoid numerical error, we compute a weighted average of left and right partial derivatives. All the numerical partial derivatives converge to zero as $n$ increases.

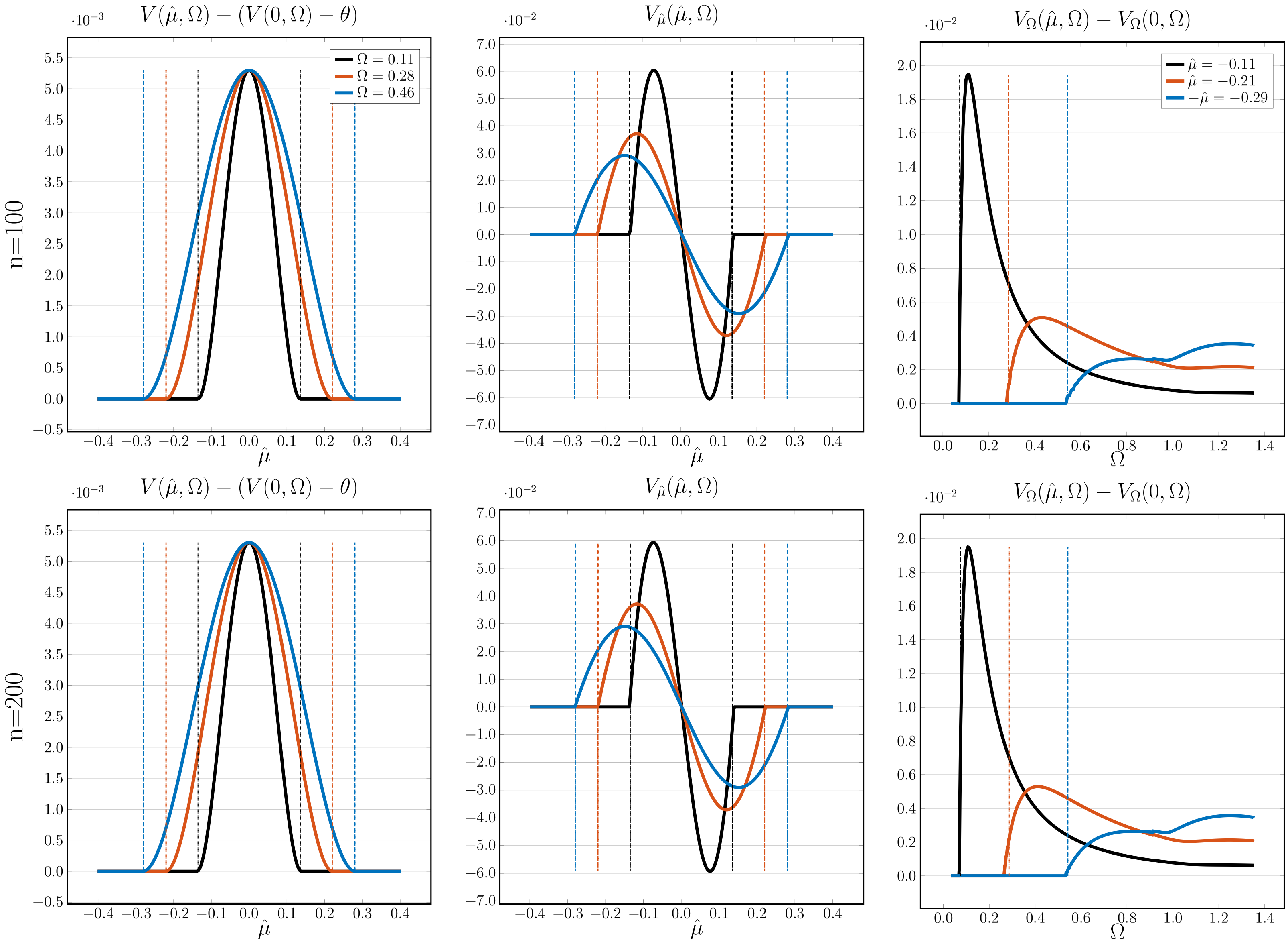
To see the results graphically, Figures V and VI plot three functions: $V(\hat{\mu}, \Omega) - (V(0, \Omega) - \theta)$, $V_{\hat{\mu}}(\hat{\mu}, \Omega)$ and $V_{\Omega}(\hat{\mu}, \Omega) - V_{\Omega}(0, \Omega)$ for different values of $n \in \{25, 50, 100, 200\}$. The first column illustrates the validity of the value matching condition. It plots the value function $V(\hat{\mu}, \Omega)$ minus the reset value $(V(0, \Omega) - \theta)$ in the markup gap space, for the three levels of uncertainty $\Omega \in \{0.11, 0.29, 0.46\}$. Outside the continuation region this difference is equal to zero, while inside the continuation region the difference is clearly positive. The second column plots the numerical partial derivative with respect to the first state $\hat{\mu}$ for the same three levels of uncertainty. We can clearly see that partial derivatives outside the continuation region are equal to zero and inside the continuation region they converge to zero as $\hat{\mu}$ gets closer to the border. Moreover, as $n$ increases, the partial derivates become smoother since the grid size and the time interval also shrink. The third column plots the numerical partial derivative with respect to $\Omega$ evaluated at the lower border of the inaction region $-\bar{\mu} \in \{-0.3, -0.21, -0.11\}$.

Figure V: Value Function and Smooth Pasting Conditions for $n = 25, 50$




First column plots the value matching condition. It plots $V(\hat{\mu},\Omega) - V(0,\Omega) - \theta)$ in the markup gap space, for the three levels of uncertainty $\Omega \in \{0.11, 0.29, 0.46\}$. Second column plots the smooth pasting condition for $\hat{\mu}$ for the same three levels of uncertainty. Third column plots the smooth pasting condition for $\Omega$ evaluated at the lower border of the inaction region $-\bar{\mu} \in \{-0.3, -0.21, -0.11\}$. The dashed lines represent the boundary of the continuation region.

Figure VI: Value Function and Smooth Pasting Conditions for $n = 100, 200$




First column plots the value matching condition. It plots $V(\hat{\mu},\Omega) - V(0,\Omega) - \theta)$ in the markup gap space, for the three levels of uncertainty $\Omega \in \{0.11, 0.29, 0.46\}$. Second column plots the smooth pasting condition for $\hat{\mu}$ for the same three levels of uncertainty. Third column plots the smooth pasting condition for $\Omega$ evaluated at the lower border of the inaction region $-\bar{\mu} \in \{-0.3, -0.21, -0.11\}$. The dashed lines represent the boundary of the continuation region.

Table III: Numerical Check of Smooth Pasting Conditions

| $n$ | $V_{\hat{\mu}}(\bar{\mu}, \Omega_1)$ | $V_{\hat{\mu}}(\bar{\mu}, \Omega_2)$ | $V_{\hat{\mu}}(\bar{\mu}, \Omega_3)$ |
|---|---|---|---|
| 25 | 0.01818 | 0.00520 | 0.00472 |
| 50 | 0.00680 | 0.00405 | 0.00192 |
| 100 | 0.00389 | 0.00217 | 0.00108 |
| 150 | 0.00340 | 0.00152 | 0.00116 |
| 200 | 0.00169 | 0.00118 | 0.00047 |
| $n$ | $V_\Omega(\bar{\mu}_1, \Omega) - V_\Omega(0, \Omega)$ | $V_\Omega(\bar{\mu}_2, \Omega) - V_\Omega(0, \Omega)$ | $V_\Omega(\bar{\mu}_3, \Omega) - V_\Omega(0, \Omega)$ |
| 25 | 0.00554 | 0.00204 | 0.00047 |
| 50 | 0.00555 | 0.00108 | 0.00046 |
| 100 | 0.00548 | 0.00059 | 0.00015 |
| 150 | 0.00222 | 0.00049 | 0.00014 |
| 200 | 0.00167 | 0.00030 | 0.00006 |

### E.4 SPC for $\Omega_t \to 0$

Let $\mathcal{D} = \{(\hat{\mu}, \Omega : |\hat{\mu}| \leq \overline{\mu}(\Omega))\}$ be the inaction region as in the main text. While we cannot ensure that the smooth pasting holds at $\Omega = 0$, be are certain that we cannot apply Theorem 2.2 Øksendal and Sulem (2010). The reason is that the probability of having a positive expected duration when the initial condition is at a boundary with zero uncertainty, i.e. $(\hat{\mu}, \Omega) = (\overline{\mu}(0), 0)$, is equal to one, and therefore the ZLM does not hold:

$$\mathbb{E}\left[\int_0^T \mathbb{1}_{\{\partial D\}}(\mu_t)\, dt \Big| (\hat{\mu}, \Omega) = (\overline{\mu}(0), 0)\right] > 0$$

To circumvent this issue, in the new version of the paper we impose that $\sigma_f > 0$ and therefore the domain of uncertainty becomes $\Omega \geq \sigma_f > 0$. Therefore, we do not need to worry about this case.

The failure of the SPC at $(\overline{\mu}(0), 0)$ can be illustrated numerically by computing the derivatives $V_\mu(\overline{\mu}(\sigma_f), \sigma_f)$ and $V_\Omega(\overline{\mu}(\sigma_f), \sigma_f) - V_\Omega(0, \sigma_f)$ and letting $\sigma_f \to 0$. If the SPC holds, the two derivatives must converge to zero. We compute the continuous time derivatives via their discrete time counterparts with an interval of size $1/n$, and then letting $n \to \infty$.

Table IV reports the values of $V_\mu(\overline{\mu}(\sigma_f), \sigma_f)$ and $V_\Omega(\overline{\mu}(\sigma_f), \sigma_f) - V_\Omega(0, \sigma_f)$ for the same calibration as in the Online Appendix, as we change $\sigma_f \in \{0.05, 0.03, 0.015, 0.007\}$.

We observe that as $\sigma \to 0$, the convergence gets slower and slower. We believe that in the limit, the SPC does not hold. We stopped at the value $\sigma = 0.007$ because of numerical limitations.

Table IV: Numerical Check of the Smooth Pasting Conditions as $\sigma_f \to 0$

| $\sigma_f = 0.05$ | $V_\mu(\overline{\mu}(\sigma_f), \sigma_f)$ | $V_\Omega(\overline{\mu}(\sigma_f), \sigma_f) - V_\Omega(0, \sigma_f)$ |
|---|---|---|
| $n$ | | |
| 25 | 0.031222591 | 0.035146856 |
| 50 | 0.034124272 | 0.028067058 |
| 100 | 0.019449471 | 0.018448208 |
| 150 | 0.013731089 | 0.013689865 |
| 200 | 0.010706146 | 0.01094095 |

| $\sigma_f = 0.03$ | $V_\mu(\overline{\mu}(\sigma_f), \sigma_f)$ | $V_\Omega(\overline{\mu}(\sigma_f), \sigma_f) - V_\Omega(0, \sigma_f)$ |
|---|---|---|
| $n$ | | |
| 25 | 0.093510512 | 0.051680683 |
| 50 | 0.038828027 | 0.046182415 |
| 100 | 0.042477751 | 0.043705967 |
| 150 | 0.024824886 | 0.029960388 |
| 200 | 0.015064022 | 0.020877406 |

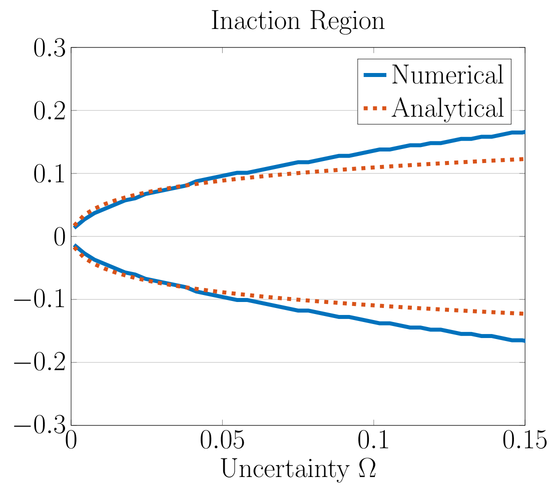
| $\sigma_f = 0.015$ | $V_\mu(\overline{\mu}(\sigma_f), \sigma_f)$ | $V_\Omega(\overline{\mu}(\sigma_f), \sigma_f) - V_\Omega(0, \sigma_f)$ |
|---|---|---|
| $n$ | | |
| 25 | 0.123042962 | 0.069711777 |
| 50 | 0.119909825 | 0.101925469 |
| 100 | 0.056650951 | 0.079162753 |
| 150 | 0.024193621 | 0.054974091 |
| 200 | 0.034310558 | 0.054951835 |

| $\sigma_f = 0.007$ | $V_\mu(\overline{\mu}(\sigma_f), \sigma_f)$ | $V_\Omega(\overline{\mu}(\sigma_f), \sigma_f) - V_\Omega(0, \sigma_f)$ |
|---|---|---|
| $n$ | | |
| 25 | 0.150226757 | 0.060828156 |
| 50 | 0.133000685 | 0.150474393 |
| 100 | 0.099869453 | 0.148578787 |
| 150 | 0.079621951 | 0.135200803 |
| 200 | 0.066961644 | 0.122585445 |

### E.5 Inaction Region

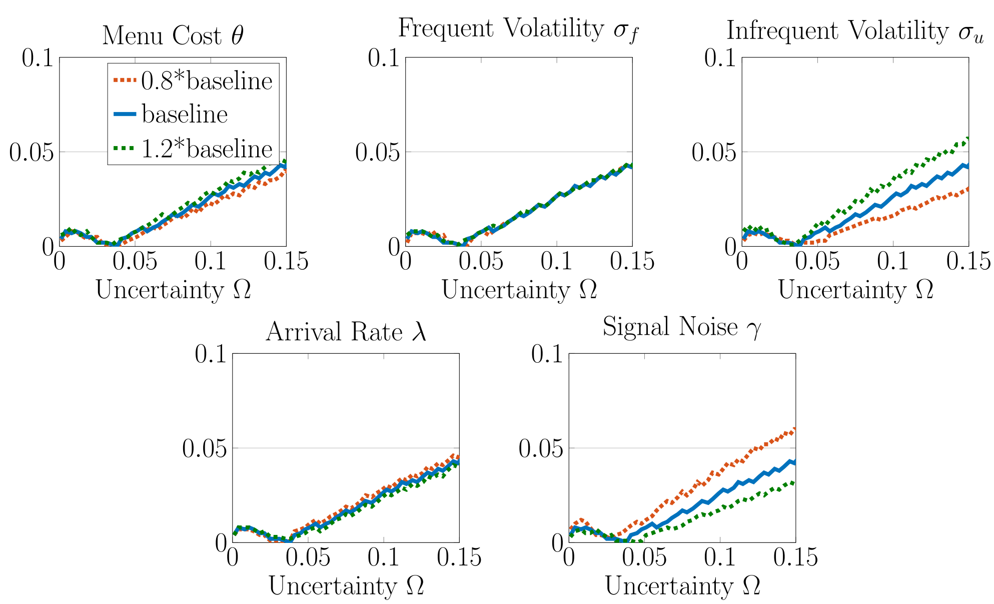
The following figure compares the inaction region $\mathcal{R} = [-\bar{\mu}(\Omega), \bar{\mu}(\Omega)]$ obtained with the analytical approximation with the one obtained by solving the continuous time model numerically. It shows the region for an uncertainty domain $\Omega \in [\sigma_f, \sigma_f + 2\frac{\sigma_u^2}{\gamma}]$. Given the calibration, 95% of firm uncertainty falls between $[\sigma_f, \sigma_f + \frac{\sigma_u^2}{\gamma}]$. We conclude that the approximation is adequate for the domain of interest.

Figure VII: Inaction Region: Numerical vs. Analytical




**Comparative statics** We vary each parameter at a time, considering half and double its benchmark value. Then we plot the approximation error defined as the absolute value of the difference between the analytical and the numerical policy. For all cases considered, the margin of error is less than 0.1.

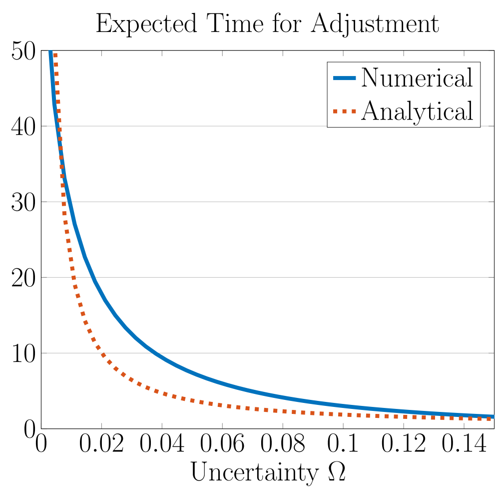
Figure VIII: Inaction Region: Approximation Error



## E.6 Expected Time

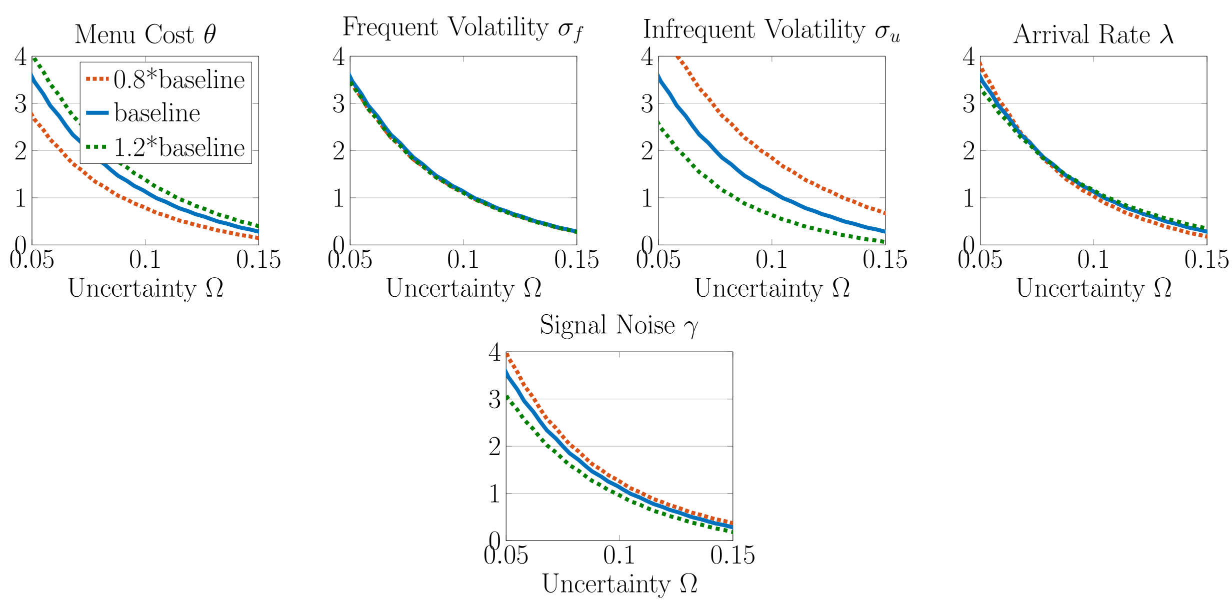
The following figure compares the expected time between price changes $\mathbb{E}[\tau|(0,\Omega)]$ obtained with the analytical approximation with the one obtained by solving the continuous time model numerically.

Figure IX: Expected Time: Numerical vs. Analytical




**Comparative statics** We now vary each parameter at a time, considering half and double its benchmark value. For each parameter, we plot the difference between the analytical and the numerical expected time. We observe that for all cases considered, the margin of error is less than 4.

Figure X: Expected Time: Approximation Error



## E.7 Hazard Rate

In all the following cases (analytical and numerical), we focus on the computation of the stopping time distribution $f(\tau)$. Given the stopping time distribution, the conditional hazard rate is computed using its definition:

$$h(\tau) \equiv \frac{f(\tau)}{\int_\tau^\infty f(s)ds}$$

**Analytical approximation with constant inaction region** Following Kolkiewicz (2002) and Álvarez, Lippi and Paciello (2011)'s Online Appendix, the density of stopping times with constant inaction region $[-\bar{\mu}_0, \bar{\mu}_0]$ is given by:

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

The process $x(\tau)$ is equal to the ratio of the state's unconditional variance to the width of the inaction region. In our case, we consider the estimate's unconditional variance which is given by $\mathcal{V}_\tau(\Omega_0)$; furthermore, we compute densities conditional on the initial level of uncertainty $\Omega_0$. For the border of inaction we set it equal to our approximation evaluated at initial uncertainty: $\bar{\mu}_0 = \bar{\mu}(\Omega_0)$. With these elements, we can apply the previous formula using $x \equiv \frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}$ and the same sequences of $\alpha_j$ and $\beta_j$ to obtain:

$$f(\tau|\Omega_0) = \frac{\pi x'(\tau)}{2} \sum_{j=0}^{\infty} (2j+1)^{-1}(-1)^j \exp\left(-(2j+1)^2 \frac{\pi^2}{8} x\right), \qquad \text{where} \quad x \equiv \frac{\mathcal{V}_\tau(\Omega_0)}{\bar{\mu}_0^2}$$

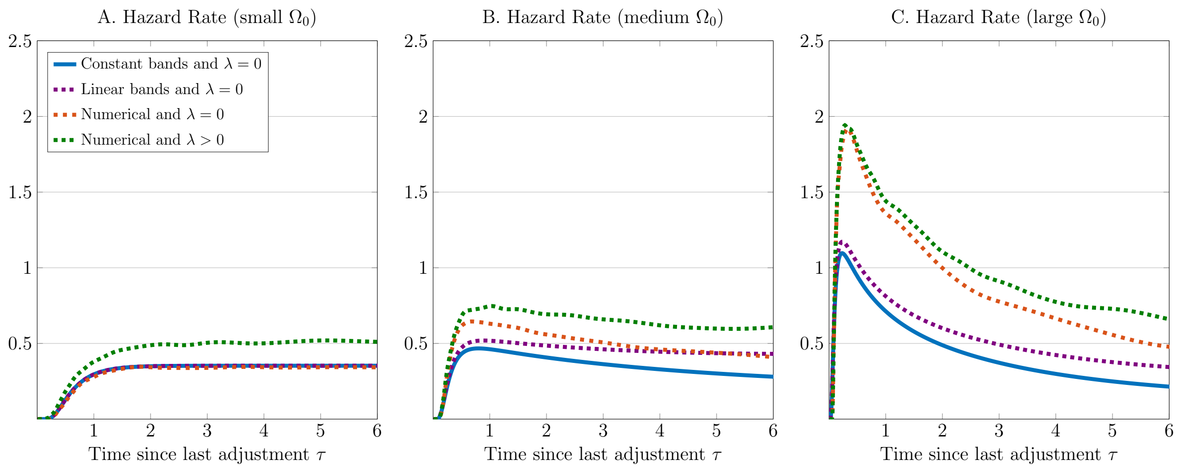
**Analytical approximation with linear inaction region** Now we assume constant inaction regions of the form $\alpha + \beta t$. The formula for the stopping time distribution is found on Theorem 3.2 of Kolkiewicz (2002). We apply the formula with values $x_0 = 0$, $b_1 = 2(\alpha + \beta t)$, $b_{2,j} = b_{3,j} = \alpha(2j+1)$ and obtain:

$$\begin{aligned} f(\tau|\Omega) &= \frac{2\alpha \mathcal{V}_\tau'(\Omega_0)}{\sqrt{\mathcal{V}_\tau(\Omega_0)^3}} \phi\left(\frac{\alpha + \beta\tau}{\sqrt{\mathcal{V}_\tau(\Omega_0)}}\right) \times \\ & \sum_{j=0}^{\infty} \left[(4j+1)\exp\left(\frac{-4j(2j+1)(\alpha+\beta\tau)\alpha}{\mathcal{V}_\tau(\Omega_0)}\right) - (4j+3)\exp\left(\frac{-4(j+1)(2j+1)(\alpha+\beta\tau)\alpha}{\mathcal{V}_\tau(\Omega_0)}\right)\right] \end{aligned} \tag{E.51}$$

where $\phi(\cdot)$ is the density of a standard Normal.

**Numerical hazard rate** To compute numerically the hazard rate, we use a Monte Carlo simulation. For this purpose, we start with a large number of firms whose state is initialized at $(0, \Omega_0)$. Then we hit firms with shocks and record the date at which a firm has her *first* price change. After her first price change, a firm is discarded. In this way, we make sure that the hazard rate we are computing is conditional on an initial level of uncertainty of $\Omega_0$ and also we do not generate right censoring. Once we have a large number of observations of stopping times, we estimate the density $f(\tau|\Omega_0)$ using a kernel smoothing function. Lastly, we compute the hazard rate using its definition: $h(\tau|\Omega_0) = \frac{f(\tau|\Omega_0)}{\int_\tau^\infty f(s|\Omega_0)ds}$.

Figure XI: Numerical Hazard Rates and Analytical Approximation




Left panel: hazard rate for $\Omega_0 = \sigma_f$. Center panel: hazard rate for $\Omega_0 = 2\sigma_f$. Right panel: hazard rate for $\Omega_0 = 5\sigma_f$. See section E.7 for the description of the calibration of the rest of the parameters.

# F Alternative Models for Uncertainty

A commonly held idea is that heterogeneity alone generates a decreasing hazard rate. *This is not the case in state dependent models, as we below with counter-examples.* The main intuition is that the hazard rate's shape depends on the dynamics of the re-pricing probability, and if the model cannot generate a large spike in this probability, followed with a decrease, then the cannot generate a decreasing hazard rate. The following are three additional counter-examples:

1. Our model of upward uncertainty cycles with signal noise $\gamma \to 0$ or $\gamma \to \infty$. In either case, the model obtains a non-decreasing hazard. See Proposition 6.

2. An AR(1) process for firm–level volatility, as in Vavra (2014), analyzed in Section F.1 below.

3. A model with "downward" uncertainty cycles analyzed in Section F.2 below.

The next two subsections describe each model, and analyze their micro and macro implications.

## F.1 Autoregressive Volatility

In this section we compare the implications of assuming an AR(1) process for (log) idiosyncratic volatility (versus assuming a Poisson process), which is a simplified version of our learning model. In both cases, we present first the continuous time environment and then the discrete time versions that we use for computational purposes.

**Autoregressive Volatility in Continuous Time** Consider a profit maximizing firm that chooses the price at which to sell its product, subject to idiosyncratic markup shocks. Let $\mu_t$ be the markup gap and $\Pi(\mu_t) = -B\mu_t^2$, $B > 0$ be the quadratic loss function. Every price change entails paying a fixed menu cost $\theta$.

The markup gap is perfectly observables, thus the information set is given by $\mathcal{I}_t = \sigma\{\mu_s : s \leq t\}$. The stochastic process for the mark gap is given by diffusion process: $d\mu_t = \Sigma_t dW_t$ where $\Sigma_t$ follows a an Ornstein-Uhlenbeck process given by

$$d\log(\Sigma_t) = -(1-\rho_v)((\log(\Sigma_t)) - \log(\sigma_f)) + \sigma_v dZ_t, \tag{F.52}$$

here $Z_t$ is a Weiner process.

**Environment in Discrete Time** Let $V(\mu, \Sigma)$ the value of a firm with current markup $\mu$ and stochastic volatility $\nu$. Analogously, let $V^{nc}(\mu, \nu)$ and $V^c(\nu)$ be the values of not changing and changing the price respectively. The the firm's recursive problem

$$\begin{aligned}
V(\mu_-, \nu_-) &= \mathbb{E}\left[\max_{c,nc}\{V^c(\nu), V^{nc}(\mu, \nu)\}\right] \\
V^{nc}(\mu, \nu) &= -B\mu^2 + \beta V(\mu, \nu) \\
V^c(\nu) &= \max_x \; \theta - Bx^2 + \beta V(x, \nu) \\
\textit{subject to} & \\
\mu &= \mu_- + \nu\epsilon \\
\log \nu &= (1-\rho_\nu)\log \sigma_f + \rho_\nu \log \nu_- + \sigma_\nu \eta \\
\eta, \epsilon &\sim \mathcal{N}(0,1).
\end{aligned}$$

For comparison, observe that the model studied in the main text has the following recursive problem over the estimates

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

**Calibration** The calibration of the baseline model is in the same as in the main text. For the AR(1) model, we use the same menu cost and discount factor. For $\nu^*, \rho_\nu, \sigma_\nu$ we match the frequency of price changes in the data, as well as the autocorrelation and cross-sectional variance of $\Sigma$ in the baseline model (with a persistence of 0.62 and a standard deviation of 0.0067 without the normalization by $\gamma$). Table V describes parameters values and compares micro-price statistics.

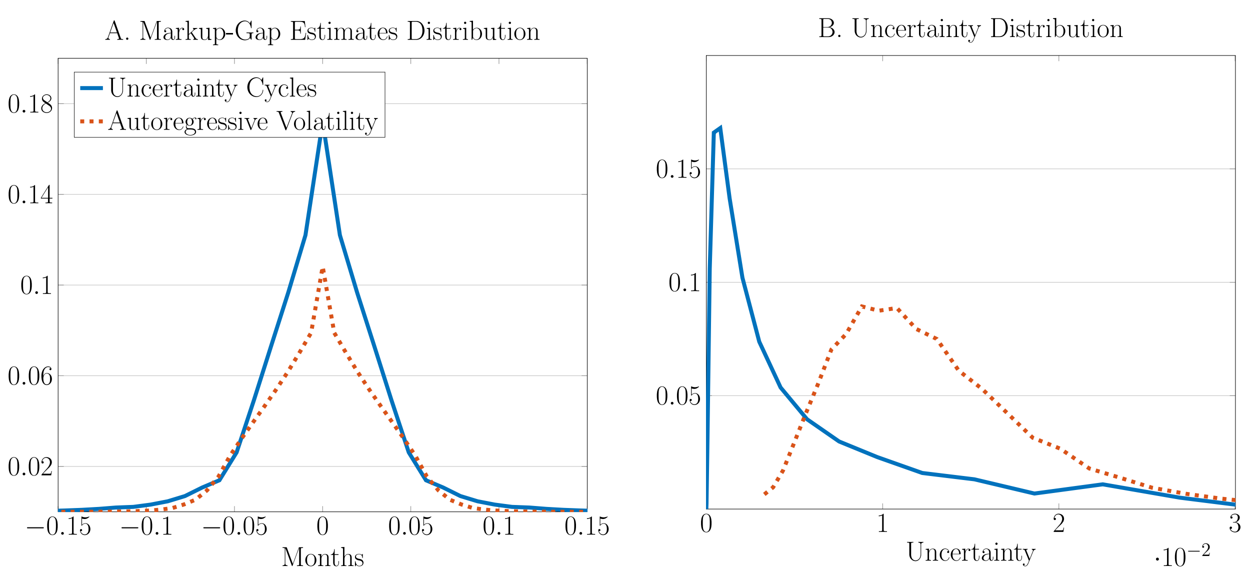
**Distribution of Uncertainty and Markup Gap (Estimates)** Figure XII plots the steady-state distributions of uncertainty and markups–gap in the stochastic volatility model and markups–gap estimates in the baseline model. The distribution of uncertainty in the baseline model is not normalized by $\gamma$. Note that the distribution generated by the BB-model is more leptokurtic (as in the data) than the AR-model. Trivially, this property implies a larger spike in the probability of price change and therefore a decreasing hazard rate.

**Hazard Rate** Figure XIII compares the re-pricing hazards for the two alternative models of heterogenous uncertainty discussed in this note. The hazard rate in the baseline model model is decreasing; the hazard rate in the AR-model is increasing then flat, and resembles the hazard rate in standard menu-cost models. The main inutition of this result, as we can see in the sample, path is that in the

Table V: Parameters and Targets

| | Uncertainty Cycles | Autoregressive Volatility |
|---|---|---|
| **Parameters** | | |
| $\sigma_f$ | 0 | 0.016 |
| $\sigma_u$ | 0.198 | 0 |
| $\lambda$ | 0.016 | 0.011 |
| $\gamma$ | 0.233 | 0.467 |
| **Moments** | | |
| $\mathbb{E}[\tau]$ in months | 10 | 10* |
| std[$\lvert\Delta p\rvert$] | 0.07 | 0.02 |
| hazard rate slope | $-0.009$ | 0.0069 |
| kurtosis[$\Delta p$] | 1.96 | 1.29 |

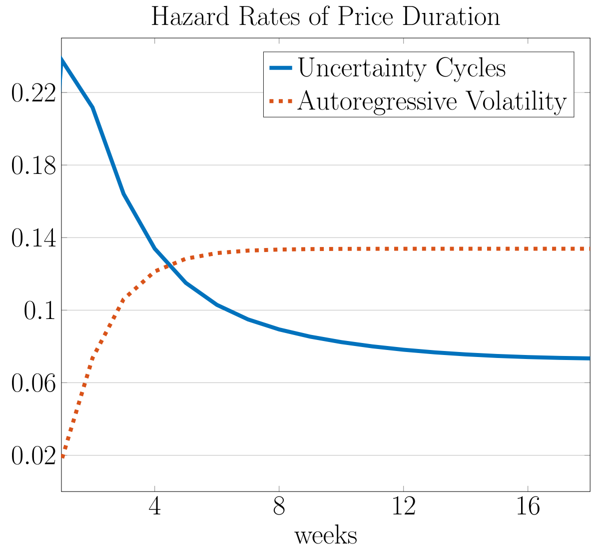
Figure XII: Marginal Distributions of Markups and Uncertainty




Panel A plots the the distribution of markup-gap estimates in both models and Panel B plots the the distribution of uncertianty in both models .

stochastic volatility model there isn't this large spikes in uncertainty follow up by a continuous and relative large decrease of uncertainty.

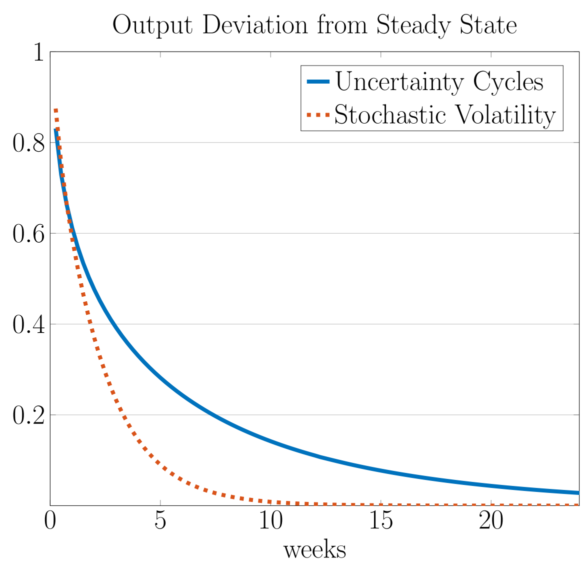
Figure XIII: Hazard Rate of Price Adjustment




This figure plots the hazard rate of price duration for the model with learning as in the main text and the model with stochastic volatility.

**Impulse-response of real output** Figure XIV shows the impulse-response of real output to an aggregate monetary shock, for the case in which this shock is perfectly disclosed ($\alpha = 1$). The BB-model amplifies the real output effects with respect to the AR-model since the two mechanism we explain in the main text are stronger: dispersion of frequency and endogenous positive correlation between selection and frequency.

Figure XIV: Output Response to a Monetary Shock



## F.2 H-model of Heterogeneity

In this subsection we explore an alternative model for generating time-varying uncertainty,[1] which we label *H-model*, in which firms perfectly learn their idiosyncratic state at random times and at other times they observe noisy signals. This model generates *downward* uncertainty cycles.

We discover that such a model produces very different predictions for micro-level price statistics, particularly the re-pricing hazard, and for the propagation of aggregate nominal shocks compared to our model in the main text, that we called BB-model. Therefore, our model of uncertainty shocks plays an important role beyond a micro-foundation for heterogeneity.

**H-model Environment in Continuous Time** Markup gaps $\mu_t$ follows a diffusion process: $d\mu_t = \sigma_f dW_t$. A Poisson process $\tilde{Q}_t$, with intensity $\lambda$, determines the dates at which the firm can perfectly observe the realization of its state. At other dates, the firm observes a noisy signal: $s_t = \mu_t dt + \gamma dZ_t$. With these assumptions, the information set is given by $\tilde{\mathcal{I}}_t = \sigma\{s_r, \tilde{Q}_r,\ r \leq t;\ \mu_k \text{ if } d\tilde{Q}_k = 1\}$. The filtering equations that describe the evolution of the markup estimate $\hat{\mu}_t \equiv \mathbb{E}[\mu_t | \tilde{\mathcal{I}}_t]$ and the uncertainty around it $\Omega_t = \frac{\mathbb{E}[(\mu_t - \hat{\mu}_t)^2 | \tilde{\mathcal{I}}_t]}{\gamma}$ are given by:

$$(estimate) \qquad d\hat{\mu}_t = \Omega_t\, d\hat{Z}_t + \gamma\, \Omega_{t\_}\, \epsilon_t\, d\tilde{Q}_t \tag{F.53}$$

$$(uncertainty) \qquad d\Omega_t = \frac{\sigma_f^2 - \Omega_t^2}{\gamma}\, dt \underbrace{-\, \Omega_{t\_}\, d\tilde{Q}_t}_{\text{downward jumps}} \tag{F.54}$$

where $d\hat{Z}_t$ is a Wiener process under $\tilde{\mathcal{I}}_t$; $\epsilon_t \sim \mathcal{N}(0,1)$ captures the difference between the current unbiased estimate and its true value; and $\Omega_{t\_} = \lim_{\Delta \to 0^+} \Omega_{t-\Delta}$ is the realization "immediately" before period $t$. Observe that (F.53) and (F.54) imply *downward* uncertainty cycles. With a "Calvo arrival of information", uncertainty jumps down to zero and then grows over time. The estimate follows a diffusion, but at those Calvo dates, it is corrected by its difference from the truth (upwards or downwards):

$$\text{if} \quad d\tilde{Q}_t = 1 \qquad \Longrightarrow \qquad \Omega_t = 0, \qquad \hat{\mu}_t = \hat{\mu}_{t\_} + \gamma \Omega_{t\_} \epsilon_t$$

With the same logic as in our model, these cycles generate heterogeneous uncertainty across a continuum of firms facing *iid* shocks.

**H-model Environment in Discrete Time** The firms' recursive problem solves

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

This recursive problem can be interpreted a time-varying noise model, where the noise is known every period.

[1] We thank Christian Hellwig for suggesting us to explore this alternative model.

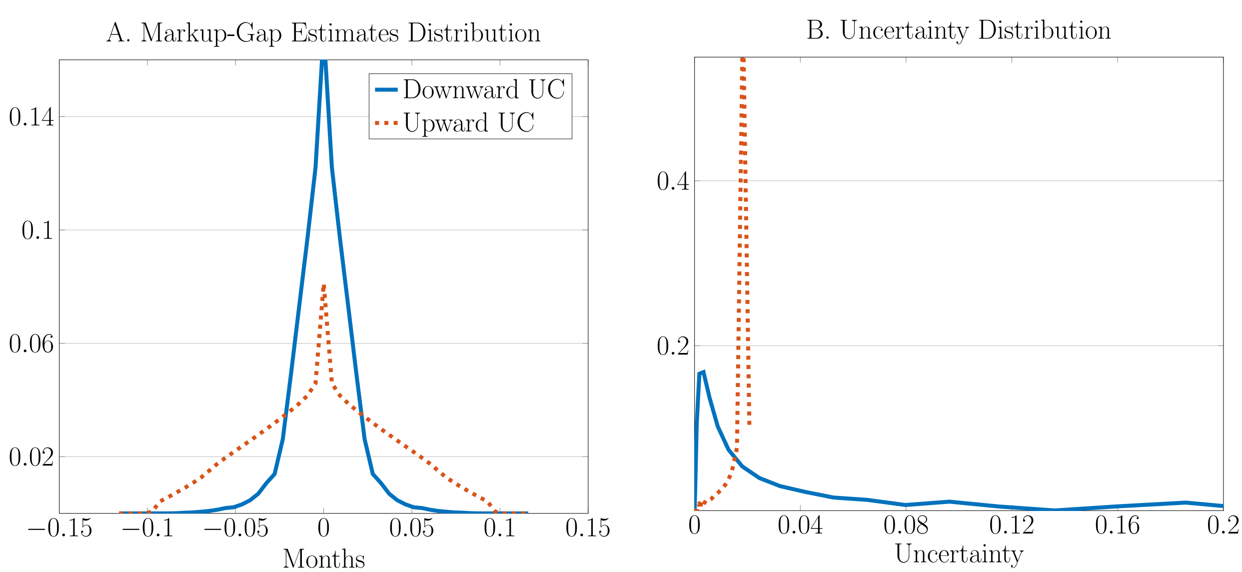
**Calibration** The common parameters across uncertainty models are $r = 0.04$, $\bar{\theta} = 0.064$, $\gamma = 0.233$ and $\lambda = 0.016$. The specific parameters in our model are $\sigma_u = 0.219$ and $\sigma_f = 0.00$, as in the main text; while in the H-model we set $\sigma_f = 0.018$ to match the expected time of a price change.

Table VI: Parameters and Targets

| | Upward Uncertainty Cycles | Downward Uncertainty Cycles |
|---|---|---|
| **Parameters** | | |
| $\sigma_f$ | 0 | 0.016 |
| $\sigma_u$ | 0.198 | 0 |
| $\lambda$ | 0.016 | 0.018 |
| $\gamma$ | 0.233 | 0.233 |
| **Moments** | | |
| $\mathbb{E}[\tau]$ in months | 10 | 10* |
| std$[\lvert\Delta p\rvert]$ | 0.05 | 0.02 |
| std$[\Delta p]$ | 0.14 | 10 |
| hazard rate slope | $-0.005$ | 0.007 |
| kurtosis$[\Delta p]$ | 1.84 | 1.17 |

**Distribution of Uncertainty and Markup Gap Estimates** Figure XV plots the steady-state distributions of uncertainty and markups–gap estimates. Note that the distribution generated by the BB-model is more leptokurtic (as in the data) than the H-model. Interestingly, both models have similar average uncertainty (0.065 in the BB-model and 0.062 in the H-model), but the BB-model generates a right-skewed distribution of uncertainty while the H-model a left-skewed distribution.

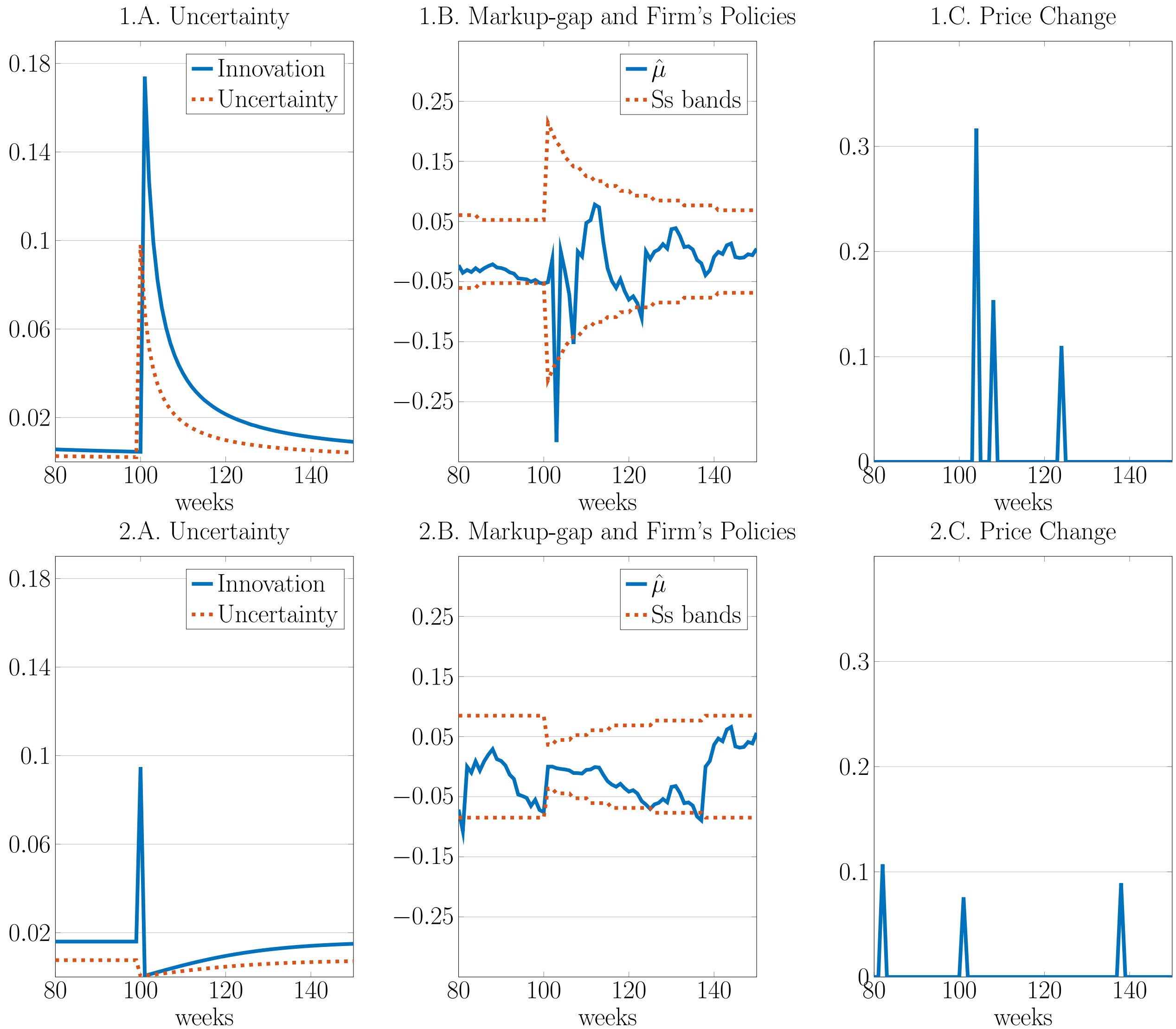
Figure XV: Marginal Distribution of the Upward and Downward Uncertainty Cycles




Panel A plots the the distribution of markup-gap estimates in both models and Panel B plots the the distribution of uncertianty in both models .

**Ilustration of Sample Paths** Figure XVI shows a sample path for a firm outcomes in each model. We plot uncertainty and the variance of markup estimates (left), inaction regions and the markup gap estimates (center), and size of price changes (right).

Figure XVI: Upward and Downward Uncertainty Cycles Dynamics



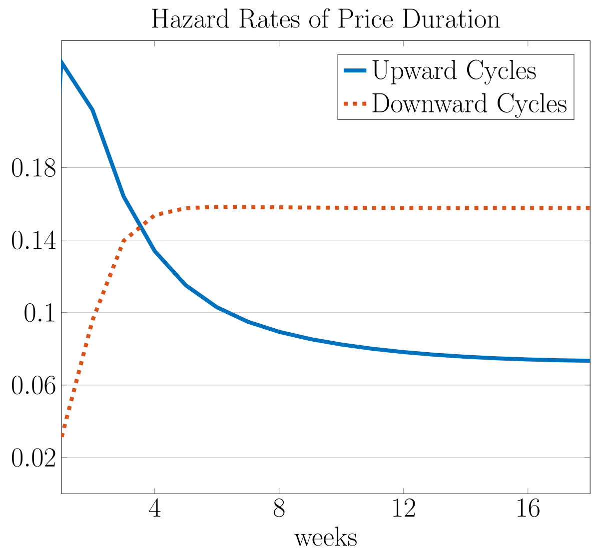

Panels 1.A to 1.C plot a sample path in the model with upward uncertainty cycles and Panels 2.A to 2.C plot a sample path in the model with downward uncertainty cycles.

**Hazard rate** Figure XVII compares the re-pricing hazards for the two alternative models of heterogenous uncertainty discussed in this note. The hazard rate in the BB-model is non-monotonic, with a decreasing segment; the hazard rate in the H-model is increasing then flat, and resembles the hazard rate in standard menu-cost models.[2]

**Impulse-response of real output** Figure XVIII shows the impulse-response of real output to an aggregate monetary shock, for the cases in which this shock is perfectly disclosed ($\alpha = 1$, Panel A)

[2]This result is robust to different set of parameters, for instance, if $\gamma \to 0$, then the model obtains a constant hazard, if $\gamma \to \infty$ or $\lambda \to 0$, the model obtains an increasing hazard.

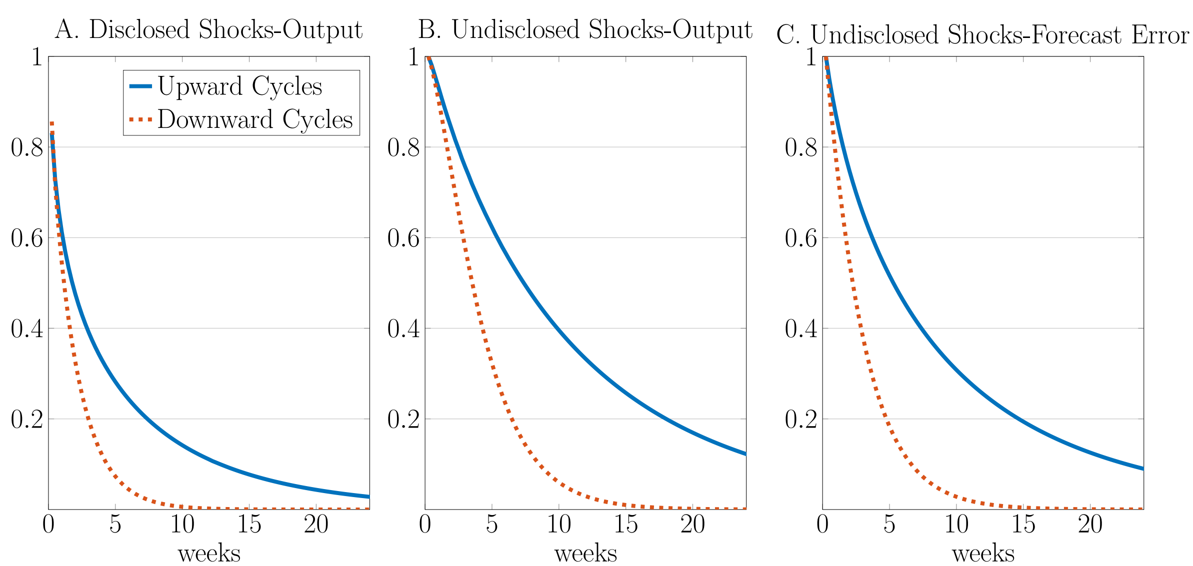
Figure XVII: Upward and Downward Uncertainty Cycles Hazard Rate




This figure plots the hazard rate of price duration for the model with downward and upward uncertainty cycles.

and not disclosed ($\alpha = 0$, Panel B). In both cases, the BB-model amplifies the real output effects with respect to the H-model. In Panel C, we plot average forecast errors for the case of undisclosed money shock; it shows that behind the larger output effect in the BB-model is the larger persistence of forecast errors.

Figure XVIII: Output Response to a Monetary Shock for the Two Types of Uncertainty Cycles




Left panel: output response for a disclosed shocks. Center panel: output response for a undisclosed shocks. Right panel: forecast error for a undisclosed shocks.

# G Firm Pricing Problem in Discrete Time

In this section, we develop a discrete time version of the firm pricing problem and show that there exists a continuous time representation when the length of the periods tends to zero. This is the continuous time problem that appears in the paper that allows us derive analytical solutions for policies and statistics.

The signal extraction problem in continuous time described in the paper—with frequent and infrequent shocks to the state together with a noisy signal—can be reinterpreted, when written in discrete time, as a problem with frequent and infrequent permanent shocks to the state together with transitory shocks, where the firm cannot distinguish between the permanent and transitory shocks. That is, the signal noise in continuous time can be reinterpreted as transitory volatility affecting the state. That is, the noise coming from the signal in the continuous time case can be thought of a transitory shock that affects the state. We prefer this second interpretation for building the economic interpretation of our model, and it is the interpretation used in the Introduction of the paper.

We will establish the link between the discrete and continuous time formulations through the filtering problem and the optimal stopping policy, this is, the time when to change the price and reset the markup process. We show that the discrete time stopping converges to the continuous time stopping as the interval length goes to zero. Although the policies converge, the values of a firm under the different measures of time do not necessarily converge.

## G.1 Environment

Time is discrete at intervals of size $\Delta$. Consider an infinitely lived firm. The per period flow profits is given by

$$\Pi(p_{t\Delta}, p^*_{t\Delta}) = -B(p_{t\Delta} - p^*_{t\Delta})^2\Delta$$

which is the distance between the currently charged log price $p_{t\Delta}$ and the target log price $p^*_{t\Delta}$. Firms choose whether to keep the price they charged last period or to change it, in which case they pay a menu cost denoted with $\theta$.

Let V be the present discounted value of profits, then the sequential problem is given by:

$$V(\cdot) = \max_{\{p_{(t)\Delta}\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} e^{-rt\Delta}\Big[-B(p_{t\Delta} - p^*_{t\Delta})^2\Delta - \theta\mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}\Big]\right] \tag{G.55}$$

where $r$ is the discount rate and $p_{t\Delta}$ is the price charged during $[t\Delta, (t+1)\Delta)$. We have omitted the firm's state because it will depend on the particular information structure assumed below.

**Observation** Notice that in the absence of menu costs, the problem becomes static and then the solution is $p_{t\Delta} = p^*_{t\Delta}$. Since prices will track the realization of the optimal price, markups are constant at all times, regardless of the stochastic process of the optimal price or the information set. As a consequence, as discussed by Hellwig and Venkateswaran (2014), there will be no output effects from a monetary shock even in the presence of information frictions.

**Productivity Process** Let $a_{t\Delta} \equiv \log(A_{t\Delta})$ denote the log of marginal productivity at $t\Delta$ and define $a^A_{t\Delta} \equiv \sum_{s=0}^{t} a_{s\Delta}\Delta$ as the accumulated productivity from 0 to $t\Delta$ periods. We can recover productivity as $a_{t\Delta} = \frac{a^A_{t\Delta} - a^A_{(t-1)\Delta}}{\Delta}$ taking $a^A_{-\Delta} = 0$. We will define the stochastic process for accumulated productivity since it will be easier to work with and it will help us map the problem into continuous time for characterization[3].

[3]See Øksendal (2007) for further details on continuous time filtering.

We assume that $a^A_{t\Delta}$ satisfies the following system of stochastic differential equations:

$$\begin{aligned} a^A_{t\Delta} &= a^A_{(t-1)\Delta} + a^P_{t\Delta}\Delta + a^T_{t\Delta} \\ a^P_{t\Delta} &= a^P_{(t-1)\Delta} + \sqrt{\Delta\sigma_f^2}\epsilon_{t\Delta} + \sqrt{\sigma_u^2}u_{t\Delta}J_{t\Delta} \\ a^T_{t\Delta} &= \sqrt{\Delta\gamma^2}\eta_{t\Delta} \\ \epsilon_{t\Delta}, \eta_{t\Delta}, u_{t\Delta} &\sim_{i.i.d} \mathcal{N}(0,1) \\ a^P_{-\Delta} &\sim \mathcal{N}(0, \Sigma_{-\Delta}) \end{aligned} \tag{G.56}$$

where $J_{t\Delta} = \begin{cases} 1 & \text{w.p. } 1 - e^{-\lambda\Delta} \\ 0 & \text{w.p. } e^{-\lambda\Delta} \end{cases}$ is an binomial random variable that takes the value of 1 with probability $1 - e^{-\lambda\Delta}$.

This formulation says that changes in accumulated productivity (current productivity) have a stochastic trend component $a^P_{t\Delta}\Delta$ and an iid transitory shock $a^T_{t\Delta}$ with variance $\gamma^2$. Innovations to the trend follow a unit root process and has two types of Normal shocks: $\varepsilon_{t\Delta}$ is a frequent shock with variance $\sigma_f^2$ and $u_t$ is a shock with variance $\sigma_u^2$ but is only received with probability $(1 - e^{-\lambda\Delta})$. The variance of the initial draw is $\Sigma_{-\Delta}$.

**Markup Gap Process** Let $\tilde{\mu}_{t\Delta} \equiv p_{t\Delta} + a_{t\Delta}$ be the log of the current markup and $\mu^* \equiv p^*_{t\Delta} + a_{t\Delta}$ the target markup. Also define $\mu_{t\Delta} \equiv \mu_{t\Delta} - \mu^*$ as the markup gap. Note that when the price is fixed, $\mu_{t\Delta}$ is driven exclusively by productivity. In accordance to the previous section, we will define the accumulated markup gap $\mu^A_{t\Delta} \equiv \sum_{s=0}^t \mu_{s\Delta}\Delta$ which allows us to recover the markup gap process as $\mu_{s\Delta} = \frac{\mu^A_{t\Delta} - \mu^A_{(t-1)\Delta}}{\Delta}$.

We use an analog decomposition into permanent and transitory shocks:

$$\begin{aligned} \mu^A_{t\Delta} &= \mu^A_{(t-1)\Delta} + \mu^P_{t\Delta}\Delta + \mu^T_{t\Delta} \\ \mu^P_{t\Delta} &= \mu^P_{(t-1)\Delta} + \sqrt{\Delta\sigma_f^2}\epsilon_{t\Delta} + \sqrt{\sigma_u^2}u_{t\Delta}J_{t\Delta} && \text{(G.57)} \\ \mu^T_{t\Delta} &= \sqrt{\Delta\gamma^2}\eta_{t\Delta} && \text{(G.58)} \\ \epsilon_{t\Delta}, \eta_{t\Delta}, u_{t\Delta} &\sim_{i.i.d} \mathcal{N}(0,1) \\ \mu^P_{-\Delta} &\sim \mathcal{N}(\mu_{-\Delta}, \Sigma_{-\Delta}) \\ J_{t\Delta} &= \begin{cases} 1 & \text{w.p. } 1 - e^{-\lambda\Delta} \\ 0 & \text{w.p. } e^{-\lambda\Delta} \end{cases} \end{aligned}$$

When the price is changed, the markup process is reset to some initial condition (that will be the optimal markup gap as shown below). The shock $\sigma_f\varepsilon^f_t$ is analogous to the continuous time frequent shocks $\sigma_f dW_t$ modeled through the Brownian motion; the shock $\sigma_u\varepsilon^U_t J_t$ is analogous to the continuous time infrequent shocks $\sigma_u u_t dQ_t$ modeled through the compound Poisson process; the shock $\gamma\varepsilon^T_t$ is analogous to the signal noise process $\gamma dZ_t$; and the binomial random variable $J_t$ with probability $1 - e^{-\lambda}$ of receiving a shock is analogous to the Poisson counting process $Q_t$ with arrival rate $\lambda$.

**Timing assumption** We assume that the firm has to set the price before the realization of the shocks in that period. We model this assumption by requiring $p_{t\Delta}$ to be $\mathcal{I}_{(t-1)\Delta}$-measurable, where $\mathcal{I}_{t\Delta}$ denotes the time $t\Delta$ information set. The objective is to eliminate the effect of the transitory shock on the pricing policy that operates through the profit function and focus only on its effect through the information structure. Under this assumption, policies will only depend on permanent shocks and we ignore price changes that respond to the transitory shocks as these are not the main point of our paper. See Kehoe and Midrigan (2015) for a paper where prices react to transitory shocks to explain discounts.

**Information Sets** We will consider two scenarios that differ in the information available for the firm at the time she makes the price adjustment decision. In the first case the firm can distinguish perfectly between permanent and transitory components of productivity. Therefore, her relevant information set at period $t\Delta$ is given by $F_{t\Delta} = \sigma\{\mu^P_{s\Delta}\Delta, \mu^T_{s\Delta}; s \leq t\}$, where $\sigma(\cdot)$ stands for the generated $\sigma$-algebra, i.e. she knows the difference between the permanent and transitory components. Under the second scenario firms have imperfect information since they cannot distinguish between the two shocks and only observe accumulated productivity. Their information set at time $t\Delta$ is given by: $G_{t\Delta} \equiv \sigma\{\mu^A_{s\Delta}, J_{s\Delta}; s \leq t\}$. We will use $\mathcal{I}_{t\Delta}$ to denote either filtration.

**A comment on the discrete time process** The previous process, with the shocks happening to the growth rate of markups, is the one that converges to the stochastic process in continuous time:

$$\begin{array}{lrcll} (State) & d\mu_t & = & \sigma_f dW_t + \sigma_u u_t dQ_t, & \mu_0 \sim \mathcal{N}(a,b) \\ (Signal) & ds_t & = & \mu_t dt + \gamma dZ_t, & s_0 = 0 \end{array}$$

where $W_t, Z_t$ are Wiener processes, $Q_t$ is a Poisson process with intensity $\lambda$, $u_t \sim \mathcal{N}(0,1)$, and $a, b$ are constants.

For building intuition, it is also convenient to show a process where the shocks occur to the level of markups and not the growth. All the results can be extended to this model. This process, although it does not converge to the continuous time process, does not change any of the ideas and results in the paper.

$$\begin{array}{rrcll} (Total) & \mu_t & = & \mu^P_t + \mu^T_t & \\ (Permanent) & \mu^P_t & = & \mu^P_{t-1} + \sigma_f \varepsilon^f_t + \sigma_u \varepsilon^u_t J_t, & \mu^P_0 = 0 \\ (Transitory) & \mu^T_t & = & \gamma \varepsilon^T_t & \\ & J_t & = & \begin{cases} 1 & \text{w.p. } 1 - e^{-\lambda} \\ 0 & \text{w.p. } e^{-\lambda} \end{cases} & \\ & \epsilon^f_t, \epsilon^u_t, \epsilon^T_t & \sim & \mathcal{N}(0,1) & \end{array}$$

## G.2 Characterization

**Expected Profits** First we convert the profit to be a function of price gaps into a function of markup gaps using that $(p_t - p^*_t) = \tilde{\mu}_{t\Delta} - a_{t\Delta} - (\mu^* - a_{t\Delta}) = \tilde{\mu}_{t\Delta} - \mu^* = \mu_{t\Delta}$:

$$\Pi(p_t - p^*_t) = \Pi(\mu_{t\Delta})$$

Now substituting the various components of the markup process we can express the profit function in terms of the underlying shocks:

$$\begin{aligned} \Pi(\mu_{t\Delta}) &= -B\mu^2_{t\Delta}\Delta \\ &= -B\left(\frac{\mu^P_{t\Delta}\Delta + \mu^T_{t\Delta}}{\Delta}\right)^2 \Delta \\ &= -B\left(\mu^P_{(t-1)\Delta} + \sigma_f\sqrt{\Delta}\epsilon_{t\Delta} + \sigma_u u_{t\Delta} J_{t\Delta} + \gamma\frac{\eta_{t\Delta}}{\sqrt{\Delta}}\right)^2 \Delta \end{aligned}$$

The expectation of profits given an information set $\mathcal{I}_t$ takes into account the timing assumption that implies orthogonality of $\epsilon_{t\Delta}, \eta_{t\Delta}$, and $u_{t\Delta}$ with respect $\mathcal{I}_{(t-1)\Delta} = \sigma\{\mu^P_{s\Delta}, \mu^T_{s\Delta} s \leq t-1\}$:

$$\mathbb{E}[\Pi(\mu_{t\Delta})|\mathcal{I}_{(t-1)\Delta}] = -B\mathbb{E}[(\mu^P_{(t-1)\Delta})^2|\mathcal{I}_{(t-1)\Delta}]\Delta - \underbrace{B\left(\sigma^2_f\Delta^2 + \sigma^2_u(1 - e^{-\lambda\Delta})\Delta + \gamma^2\right)}_{\text{sunk costs}}$$

Observe that expected profits depend exclusively on the permanent component of markup gap (or its expectation) and all the volatilities enter as sunk costs and will not affect the firm's policy through the profit channel, only through the information structure.

**Stopping time representation** The following lemmas restate the problem in stopping time representation, which will be used to cast the problem in continuous time. Lemma 1 in particular shows that with perfect information the firm only cares about permanent shocks as transitory shocks enter as a sunk costs and do not affect her policies via the profit channel.

**Lemma 1** ***(Stopping time with perfect information)*** *Let $V(\mu^P_{-\Delta})$ be the value of a firm with perfect information and an initial permanent markup gap $\mu^P_{\Delta}$. Let $\tau_\Delta$ denote the optimal stopping time policy. Then $V(\mu^P_{-\Delta})$ is the solution to*

$$V(\mu^P_{-\Delta}) = \max_{\tau_\Delta} \mathbb{E}\left[\sum_{t=0}^{\tau_\Delta} -e^{-rt\Delta} B\left((\mu^P_{(t-1)\Delta})^2\Delta + \Delta^2\sigma_f^2 + \sigma_u^2(1-e^{-\lambda\Delta}) + \gamma^2\right) + e^{-r\tau_\Delta\Delta}\left[-\theta + \max_x V(x)\right]\right] \tag{G.59}$$

*with*

$$\mu^P_{t\Delta} = \mu^P_{(t-1)\Delta} + \sigma_f\sqrt{\Delta}\epsilon_{t\Delta} + \sigma_u u_{t\Delta} J_{t\Delta}$$

**Proof.** From the sequential problem in (G.55) we have

$$V(\mu^P_{-\Delta}) = \max_{\{p_{t\Delta}\}_0^\infty} \mathbb{E}\left[\sum_{t=0}^{\infty} e^{-rt\Delta}[\Pi(\mu_{t\Delta}) - \theta\mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}\right]$$

By the law of iterated expectations[4] and given that $\mu_{t\Delta}$ is measurable with respect to $F_{(t-1)\Delta} = \sigma\{\mu^P_{s\Delta}\Delta, \mu^T_{s\Delta}; s \leq t\}$:

$$\begin{aligned}
V(\mu^P_{-\Delta}) &= \max_{\{\mu_{t\Delta}\}_{t=0}^\infty} \sum_{t=0}^{\infty} e^{-rt\Delta}\mathbb{E}\left[\Pi(\mu_{t\Delta})\Delta - \theta\mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}\right] \\
&= \max_{\{\mu_{t\Delta}\}_{t=0}^\infty} \sum_{t=0}^{\infty} e^{-rt\Delta}\mathbb{E}\left[\mathbb{E}[\Pi(\mu_{t\Delta})\Delta - \theta\mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}|F_{(t-1)\Delta}]\right] \\
&= \max_{\{\mu_{t\Delta}\}_{t=0}^\infty} \sum_{t=0}^{\infty} e^{-rt\Delta}\mathbb{E}\left[\mathbb{E}[\Pi(\mu_{t\Delta})\Delta|F_{(t-1)\Delta}] - \theta\mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}\right] \\
&= \max_{\{\mu_{t\Delta}\}_{t=0}^\infty} \sum_{t=0}^{\infty} e^{-rt\Delta}\mathbb{E}\left[-B\left((\mu^P_{(t-1)\Delta})^2\Delta + \sigma_f^2\Delta^2 + \sigma_u^2(1-e^{-\lambda\Delta})\Delta + \gamma^2\right) - \theta\mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}\right]
\end{aligned}$$

By the Principle of Optimality

$$V(\mu^P_{-\Delta}) = \max_{\tau_\Delta} \mathbb{E}\left[\sum_{t=0}^{\tau_\Delta} -e^{-rt\Delta} B\left((\mu^P_{(t-1)\Delta})^2\Delta + \Delta^2\sigma_f^2 + \sigma_u^2(1-e^{-\lambda\Delta}) + \gamma^2\right) + e^{-r\tau_\Delta\Delta}\Big[-\theta + \max_x V(x)\Big]\right]$$

where permanent shocks follow

$$\mu^P_{t\Delta} = \mu^P_{(t-1)\Delta} + \sigma_f\sqrt{\Delta}\epsilon_{t\Delta} + \sigma_u u_{t\Delta} J_{t\Delta}$$

■

[4] We also apply the Monotone Convergence Theorem to exchange the expectation and infinite sum operators since each random variable is non-negative.

Now consider the case with imperfect information. The firm will forecast the permanent markup gap using the Kalman Filter to extract information from the signal given by the accumulated markup gap. Let $\hat{\mu}^P_{t\Delta} = \mathbb{E}[\mu^P_{t\Delta}|G_{(t-1)\Delta}]$ the forecast of the permanent markup conditional on the information set $G_{(t-1)\Delta} \equiv \sigma\{\mu^A_{s\Delta}, J_s; s \le t\}$ and $\Sigma_{t\Delta} = \mathbb{E}[(\hat{\mu}^P_{t\Delta} - \mu^P_{t\Delta})(\hat{\mu}^P_{t\Delta} - \mu^P_{t\Delta})'|G_{(t-1)\Delta}]$ the variance of such forecast. The state of the firm in this case is $(\hat{\mu}^P_{t\Delta}, \Sigma_{t\Delta})$.

**Lemma 2** ***(Stopping time with imperfect information)*** *Let $V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta})$ be the value of a firm with imperfect information, initial forecast $\hat{\mu}^P_{-\Delta}$ and forecast variance $\Sigma_{-\Delta}$. Let $\tau_\Delta$ denote the optimal stopping time policy. Then $V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta})$ solves*

$$V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) = \max_{\tau_\Delta} \mathbb{E}\left[\sum_{t=0}^{\tau_\Delta} -e^{-rt\Delta} B\left((\hat{\mu}^P_{t\Delta})^2\Delta + \Sigma_{t\Delta}\Delta + \gamma^2\right) + e^{-r\tau_\Delta\Delta}\Big[-\theta \max_x V(x, \Sigma_{(\tau-1)\Delta})\Big]\right] \tag{G.60}$$

*subject to the forecasting equations of the Kalman Filter:*

$$\begin{aligned}
\hat{\mu}^P_{t\Delta} &= \hat{\mu}^P_{(t-1)\Delta} + K_{t\Delta}[\mu_{(t-1)\Delta} - \hat{\mu}^P_{(t-1)\Delta}] \\
\Sigma_{t\Delta} &= (1-K_{t\Delta})^2\Sigma_{(t-1)\Delta} + K^2_{t\Delta}\frac{\gamma^2}{\Delta} + \sigma_f^2\Delta + \sigma_u^2 J_{t\Delta} \qquad \text{(G.61)} \\
K_{t\Delta} &= \frac{\Sigma_{(t-1)\Delta}}{\Sigma_{(t-1)\Delta} + \frac{\gamma^2}{\Delta}} \\
(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) & \qquad \textit{Given} \qquad \text{(G.62)}
\end{aligned}$$

**Proof.** Let $V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta})$ be the value of a firm with imperfect information. We apply the Law of Iterated Expectations to condition on $G_{(t-1)\Delta} = \sigma\{\mu^A_{s\Delta}, s \le t-1\}$ and use the definition of the profit function recalling that $\mu_{(t-1)\Delta} = \frac{\mu^A_{(t-1)\Delta} - \mu^A_{(t-2)\Delta}}{\Delta}$ to get:

$$\begin{aligned}
V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) &= \max_{\{\mu_{t\Delta}\}_{t=0}^\infty} \sum_{t=0}^{\infty} e^{-rt\Delta}\mathbb{E}\left[\mathbb{E}[\Pi(\mu_{t\Delta})|G_{(t-1)\Delta}]\Delta - \theta \mathbb{1}_{\{p_{t\Delta} \neq p_{(t-1)\Delta}\}}\right] \\
&= \max_{\{\mu_{t\Delta}\}_{t=0}^\infty} \sum_{t=0}^{\infty} e^{-rt\Delta}\mathbb{E}\left[-B\mathbb{E}\left[\mu^2_{t\Delta}\middle|\mu_{(t-1)\Delta}\right]\Delta - \theta \mathbb{1}_{\{p_{t\Delta} \neq p_{(t-1)\Delta}\}}\right] \qquad \text{(G.63)}
\end{aligned}$$

In order to work out the expression for $\mathbb{E}\left[\mu^2_{t\Delta}\middle|\mu_{(t-1)\Delta}\right]$, we need to establish the laws of motion for the forecasts. The state-space representation of the hidden Markov process $\mu^P_{t\Delta}$ and of its noisy signals $\mu_{t\Delta} = \frac{\mu^A_{t\Delta} - \mu^A_{(t-1)\Delta}}{\Delta}$ is the following:

$$\begin{aligned}
\mu^P_{t\Delta} &= \mu^P_{(t-1)\Delta} + \sigma_f\sqrt{\Delta}\epsilon_{t\Delta} + \sigma_u u_{t\Delta}J_{t\Delta} \\
\mu_{t\Delta} &= \mu^P_{t\Delta} + \sqrt{\frac{\gamma^2}{\Delta}}\eta_{t\Delta} \\
\mu^P_{-\Delta} &\sim N(\mu_{-\Delta}, \Sigma_{-\Delta}) \qquad \text{(G.64)} \\
\epsilon_{t\Delta}, \eta_{t\Delta}, u_{t\Delta} &\sim \mathcal{N}(0,1)
\end{aligned}$$

$$J_{t\Delta} = \begin{cases} 1 & \text{w.p. } 1 - e^{-\lambda\Delta} \\ 0 & \text{w.p. } e^{-\lambda\Delta} \end{cases}$$

Let $\hat{\mu}^P_{t\Delta} = \mathbb{E}[\mu^P_{t\Delta}|\mu_{(t-1)\Delta}]$ the conditional forecast and $\Sigma_{t\Delta} = \mathbb{E}[(\hat{\mu}^P_{t\Delta} - \mu^P_{t\Delta})(\hat{\mu}^P_{t\Delta} - \mu^P_{t\Delta})'|\mu_{(t-1)\Delta}]$ the conditional variance of such forecast. Applying the Kalman Filter to the previous system yields the forecasting equations:

$$\begin{aligned}\hat{\mu}^P_{t\Delta} &= \hat{\mu}^P_{(t-1)\Delta} + K_{t\Delta}[\mu_{(t-1)\Delta} - \hat{\mu}^P_{(t-1)\Delta}] \\ \Sigma_{t\Delta} &= (1-K_{t\Delta})^2\Sigma_{(t-1)\Delta} + K^2_{t\Delta}\frac{\gamma^2}{\Delta} + \sigma_f^2\Delta + \sigma_u^2 J_{t\Delta} \\ K_{t\Delta} &= \frac{\Sigma_{(t-1)\Delta}}{\Sigma_{(t-1)\Delta} + \frac{\gamma^2}{\Delta}}\end{aligned} \tag{G.65}$$

$$(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) \qquad \text{Given} \tag{G.66}$$

where $K_{t\Delta}$ is the Kalman gain. Using the previous equations on the expression of interest:

$$\begin{aligned}\mathbb{E}[\mu^2_{t\Delta}|\mu_{(t-1)\Delta}] &= \mathbb{E}[\mu_{t\Delta}|\mu_{(t-1)\Delta}]^2 + \mathbb{V}[\mu_{t\Delta}|\mu_{(t-1)\Delta}] \\ &= (\hat{\mu}^P_{t\Delta})^2 + \Sigma_{t\Delta} + \frac{\gamma^2}{\Delta}\end{aligned} \tag{G.67}$$

From (G.63) and (G.67) we have that

$$V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) = \max_{\{\mu_{t\Delta}\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} e^{-rt\Delta}\left(-B\left((\hat{\mu}^P_{t\Delta})^2\Delta + \Sigma_{t\Delta}\Delta + \gamma^2\right) - \theta \mathbb{1}_{\{p_{t\Delta}\neq p_{(t-1)\Delta}\}}\right)\right]$$

Finally, from the Principle of Optimality

$$V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) = \max_{\tau_\Delta} \mathbb{E}\left[\sum_{t=0}^{\tau_\Delta} -e^{-rt\Delta}B\left((\hat{\mu}^P_{t\Delta})^2\Delta + \Sigma_{t\Delta}\Delta + \gamma^2\right) + e^{-r\tau_\Delta\Delta}\Big[-\theta + \max_x V(x, \Sigma_{(\tau-1)_\Delta})\Big]\right]$$

subject to (G.65).

When the price is changed, the markup process is reset to $\hat{\mu}^P_{-\Delta} \sim \mathcal{N}(x, \Sigma_{(\tau-1)\Delta})$, where $x$ is the argmax above. Note that the forecast variance $\Sigma_{t\Delta}$ is not reset and it keeps following its path. ∎

By using the forecast of permanent markup instead of the actual value, firms with imperfect information have larger uncertainty coming from the forecasting process. Since the markup gap is Normally distributed, only the first two moments of the distribution are necessary for Bayesian updating and the estimate provided by the Kalman Filter is optimal. Lastly, note that the innovations are white noise.

### G.3 Continuous Time Limit of Discrete Problem

In this section, we show that there exists a continuous time representation of the problems above when the length of the periods tends to zero. This is the continuous time problem which appears in the paper that allows us derive analytical solutions for policies and statistics.

**Proposition 4** ***(Convergence of discrete time policy with perfect information)*** *Let $\tau_\Delta\Delta$ be the optimal stopping policy of the perfect information given by (G.59). Then $\tau_\Delta\Delta \to \tau$ where $\tau$ solves*

$$V(\mu^P_0) = \max_{\tau} \mathbb{E}_0\left[\int_0^{\tau} -e^{-rt}B(\mu^P_t)^2 dt + e^{-r\tau}\Big[-\theta + \max_x V(x)\Big]\right] \tag{G.68}$$

*subject to the stochastic process*

$$d\mu^P_t = \sigma_f dF_t + \sigma_u \epsilon_t dQ_t, \qquad \mu^P_0 \text{ given}$$

*where $F_t$ a standard Brownian Motion, $Q_t$ is a Poisson process with arrival rate $\lambda$ and $\varepsilon_t \sim \mathcal{N}(0,1)$. The process is reset to $\mu^P_0 = x$ with every price change.*

**Proposition 5** ***(Convergence of discrete time policy with imperfect information)*** *Let $\tau_\Delta \Delta$ be the optimal stopping policy of the imperfect information problem given by (G.60). Then $\tau_\Delta \Delta \to \tau$ where $\tau$ solves*

$$V(\hat{\mu}_0, \Omega_0) = \max_{\tau} \mathbb{E}\left[\int_0^\tau -e^{-rt} B \hat{\mu}_t{}^2 dt + e^{-r\tau}\Big[-\theta + \max_x V(x, \Omega_\tau)\Big]\right] \tag{G.69}$$

*subject to the stochastic processes*

$$\begin{aligned} d\hat{\mu}_t &= \Omega_t dW_t \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(G.70)} \\ d\Omega_t &= \frac{1}{\gamma}\Big[\sigma_f^2 - \Omega_t^2\Big]dt + \frac{\sigma_u^2}{\gamma} dQ_t \quad \textit{with} \quad \Omega_0 = \Sigma_0/\gamma \textit{ given} \end{aligned}$$

*where $W_t$ is a standard Brownian motion and $Q_t$ is a Poisson process with arrival rate $\lambda$. $\Omega_t = \frac{\Sigma_t}{\gamma}$ is a normalized forecast variance. When the price is changed, the process for markup gaps is reset to $\hat{\mu}_0 = x$ but $\Omega_t$ is not reset, it continues its path given by (G.70).*

**Proof.** (**Proposition 1 and 2**) We give the proof for Proposition 2, for 1 is analogous. Start with the discrete time problem in stopping time representation (G.60) from Lemma 2:

$$\begin{aligned} V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) &= \max_{\tau_\Delta} \mathbb{E}\left[\sum_{t=0}^{\tau_\Delta} -e^{-rt\Delta} B\left((\hat{\mu}^P_{t\Delta})^2\Delta + \Sigma_{t\Delta}\Delta + \gamma^2\right) + e^{-r\tau_\Delta\Delta}\Big[-\theta \max_x V(x, \Sigma_{(\tau-1)_\Delta})\Big]\right] \\ \hat{\mu}^P_{t\Delta} &= \hat{\mu}^P_{(t-1)\Delta} + K_{t\Delta}[\mu_{(t-1)\Delta} - \hat{\mu}^P_{(t-1)\Delta}] \qquad\qquad \text{(G.71)} \\ \Sigma_{t\Delta} &= (1-K_{t\Delta})^2\Sigma_{(t-1)\Delta} + K^2_{t\Delta}\frac{\gamma^2}{\Delta} + \sigma_f^2\Delta + \sigma_u^2 J_{t\Delta} \qquad\qquad \text{(G.72)} \\ K_{t\Delta} &= \frac{\Sigma_{(t-1)\Delta}}{\Sigma_{(t-1)\Delta} + \frac{\gamma^2}{\Delta}} \\ (\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) & \quad \text{given} \end{aligned}$$

Now we proceed in three steps:

a) Write the value function without sunk costs

Define $v_\Delta$ and $W_\Delta$ as follows:

$$\begin{aligned} v_\Delta(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) &\equiv \max_{\tau_\Delta} \mathbb{E}\left[\sum_{t=0}^{\tau_\Delta} -Be^{-rt\Delta}(\hat{\mu}^P_{(t-1)\Delta})^2\Delta + e^{-r\tau_\Delta\Delta}[-\theta + \max_x v(x, \Sigma_{(\tau-1)\Delta})\right] \\ W_\Delta(\Sigma_{-\Delta}) &\equiv \sum_{t=0}^{\infty} -Be^{-rt\Delta}[\Sigma_{(t-1)\Delta}\Delta + \gamma^2] \end{aligned}$$

By standard arguments, we rewrite the original value function as:

$$V(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) = v_\Delta(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta}) + W_\Delta(\Sigma_{-\Delta})$$

Since the stopping time can be characterized with $v_\Delta(\hat{\mu}^P_{-\Delta}, \Sigma_{-\Delta})$, we will work directly with this formulation.

b) Local consistency of stochastic processes

Given that $\mu_{(t-1)\Delta} - \hat{\mu}^P_{(t-1)\Delta} \sim N(0, \Sigma_{(t-1)\Delta} + \frac{\gamma^2}{\Delta})$, we can rewrite (G.71) and (G.72) as

$$\begin{aligned}
\hat{\mu}^P_{t\Delta} &= \hat{\mu}^P_{(t-1)\Delta} + \Sigma_{(t-1)\Delta}\sqrt{\frac{\Delta}{\Delta\Sigma_{(t-1)\Delta} + \gamma^2}}\epsilon_{t\Delta} \\
\Sigma_{t\Delta} &= \Sigma_{(t-1)\Delta} + \sigma_f^2\Delta + \sigma_u^2 J_{t\Delta} + \frac{\Delta\Sigma^2_{(t-1)\Delta}\gamma^2}{(\Delta\Sigma_{(t-1)\Delta} + \gamma^2)^2} - \frac{2\Sigma^2_{(t-1)\Delta}\Delta}{\Delta\Sigma_{(t-1)\Delta} + \gamma^2} + \frac{\Delta^2\Sigma^3_{(t-1)\Delta}}{(\Delta\Sigma_{(t-1)\Delta} + \gamma^2)^2}
\end{aligned}$$

Or equivalently

$$\begin{aligned}
\hat{\mu}^P_{t\Delta} &= \hat{\mu}^P_{(t-1)\Delta} + \Sigma_{(t-1)\Delta}\varphi^I(\Delta)\epsilon_{t\Delta} \\
\Sigma_{t\Delta} &= \Sigma_{(t-1)\Delta} + (\sigma_f^2 - \varphi^{II}(\Delta))\Delta + \sigma_u^2 J_{t\Delta} + \varphi^{III}(\Delta)
\end{aligned}$$

Where the $\varphi$ functions have the following limits

$$\begin{aligned}
\varphi^I(\Delta) &= \sqrt{\frac{\Delta}{\Delta\Sigma_{(t-1)\Delta} + \gamma^2}} && \text{with } \frac{\varphi^I(\Delta)}{\frac{\sqrt{\Delta}}{\gamma}} \to_{\Delta\to 0} 1 \\
\varphi^{II}(\Delta) &= \frac{2\Sigma^2_{(t-1)\Delta}}{\Delta\Sigma_{(t-1)\Delta} + \gamma^2} - \frac{\Sigma^2_{(t-1)\Delta}\gamma^2}{(\Delta\Sigma_{(t-1)\Delta} + \gamma^2)^2} && \text{with } \frac{\varphi^{II}(\Delta)}{\frac{\Sigma^2_{(t-1)\Delta}}{\gamma^2}} \to_{\Delta\to 0} 1 \\
\varphi^{III}(\Delta) &= \frac{\Delta^2\Sigma^3_{(t-1)\Delta}}{(\Delta\Sigma_{(t-1)\Delta} + \gamma^2)^2} && \text{with } \varphi^{III}(\Delta) = o(\Delta)
\end{aligned}$$

Now, note that with probability $1 - e^{-\lambda\Delta}$

$$\mathbb{E}[\hat{\mu}^P_{t\Delta} - \hat{\mu}^P_{(t-1)\Delta}] = 0; \qquad \mathbb{E}[(\hat{\mu}^P_{t\Delta} - \hat{\mu}^P_{(t-1)\Delta})^2] = \frac{\Sigma^2_{(t-1)\Delta}}{\gamma^2}\Delta + o(\Delta)$$

$$\mathbb{E}[\Sigma_{t\Delta} - \Sigma_{(t-1)\Delta}] = \Delta\left(\sigma_f^2 - \frac{\Sigma^2_{(t-1)\Delta}}{\gamma^2}\right) + o(\Delta); \qquad \mathbb{E}[(\Sigma_{t\Delta} - \Sigma_{(t-1)\Delta})^2] = o(\Delta)$$

and with probability $e^{-\lambda\Delta}$

$$(\hat{\mu}^P_{t\Delta}, \Sigma_{t\Delta} | (\hat{\mu}^P_{(t-1)\Delta}, \Sigma_{(t-1)\Delta}) = (\hat{\mu}^P_{(t-1)\Delta}, \Sigma_{(t-1)\Delta} + \sigma_u^2)$$

Therefore the process is locally consistent with the process

$$\begin{aligned}
d\hat{\mu}^P_t &= \frac{\Sigma_t}{\gamma}dW_t \\
d\Sigma_t &= (\sigma_f^2 - \frac{\Sigma_t^2}{\gamma^2})dt + \sigma_u^2 dQ_t
\end{aligned}$$

where $W_t$ is a Brownian motion and $Q_t$ is a Poisson process with arrival late $\lambda$.

c) Apply Convergence Theorem

Let $v$ and $\tau$ satisfy the following Bellman equation

$$\begin{aligned}
v(\hat{\mu}^P_0, \Sigma_0) &= \max_\tau \mathbb{E}\left[\int_{t=0}^{\tau} -Be^{-rt}(\hat{\mu}^P_t)^2 dt + e^{-r\tau}[-\theta + \max_x v(x, \Sigma_\tau)]\right] \\
d\hat{\mu}^P_t &= \Omega_t dW_t \\
d\Omega_t &= \frac{\sigma_f^2 - \Omega_t^2}{\gamma}dt + \frac{\sigma_u^2}{\gamma}dQ_t
\end{aligned}$$

Then by Theorem 2.1 in chapter 10 in Kushner and Dupuis (2013) we have that

$$(v_\Delta, \tau_\Delta) \rightarrow_{weakly} (v, \tau)$$

■

# H Uncertainty cycles vs. fixed heterogeneity

In this section we discuss the differences between fixed heterogeneity and uncertainty cycles. With fixed heterogeneity, there is constant but different uncertainty across firms. In our model with uncertainty cycles, uncertainty is time-varying at the firm-level, but all firms are ex-ante identical. To analyze the different implications, we set up an environment analogous to ours but with fixed heterogeneity. In this environment we show there are two key difference between fixed heterogeneity and time varying heterogeneity: (1) Time-varying heterogeneity allows us to identify the variance of the noise shocks; (2) Time-varying heterogeneity generates an endogenous negative covariance between the expected time between price changes and the selection effect embedded in price changes, and we show that this covariance increases the effect of a monetary shock.

## H.1 Two alternative models

Our model with ex-ante identical firms and time varying uncertainty we have 5 parameters: the normalized menu cost given $\theta$, the variance of the Brownian process for persistence innovations in markup-shocks given by $\sigma_f$, the variance and the arrival of the compound Poisson process for the persistence innovations of the markup-shocks given by $(\lambda, \sigma_u)$, and the variance of the transitory shocks given by $\gamma$.

In the model with fixed uncertainty, each firm has a variance $\sigma_{if}$ and we shut down the Poisson process $\sigma_{ui} = 0$. The following table describes the set of parameters in each model:

Table VII: Parameter Space in Each Model

| Time-Varying and Ex-ante Homogeneous Uncertainty | Fixed and Heterogeneous Uncertainty |
|---|---|
| Menu Cost ($\theta$) | Menu Cost ($\theta$) |
| Brownian permanent innovation ($\sigma_f$) | Brownian Permanent innovation ($\sigma_{if}$) |
| Poisson permanent innovation $(\lambda, \sigma_u)$ | - |
| Brownian transitory innovation ($\gamma$) | Brownian transitory innovation ($\gamma$) |

## H.2 Identification of noise shock volatility

In the main text we show that our model is able to identify $\gamma$ using micro-price statistics, in particular the hazard rate. The model with fixed and heterogeneous uncertainty cannot. The proof is trivial since the process for the mark-up gap estimate is independent of $\gamma$ and given by $d\hat{\mu}_t = \sigma_i dZ_t$, thus micro-price statistics do not depend on $\gamma$.

## H.3 Endogenous covariance between expected time and selection effects

We will characterize the effect of a monetary shock in the model with fixed uncertainty. Although we will only assume heterogeneity in uncertainty, our results are valid for heterogeneity in all parameters. Since each firm has different uncertainty the expected time and the kurtosis of price changes are different across firms. Let $\mathcal{K}_i, \mathcal{T}_i$ denote the idiosyncratic kurtosis and expected time, and $\mathcal{K}, \mathcal{T}$ the aggregate kurtosis of price change and expected time.

**Proposition** *In the model with fixed and heterogeneous uncertainty, the effect of a disclosed small money shock is given by*

$$\frac{\mathcal{M}(\delta)}{\delta} = \frac{\mathbb{E}[\mathcal{K}_i]\mathbb{E}[\mathcal{T}_i] + \mathbb{C}ov[\mathcal{K}_i, \mathcal{T}_i]]}{6} = \frac{\mathbb{E}_i[\sigma_i^2]^2\mathcal{K}\mathcal{T} - \mathbb{C}ov(\mathcal{K}_i\mathcal{T}_i, \sigma_i^4)}{6\mathbb{E}_i[\sigma_i^4]} \tag{H.73}$$

**Proof.** To show this result we proceed in 5 steps.

1. *Calculation of the Renewal Distribution:* In each period of time $\Delta$, a measure $\Delta\mathcal{T}_i^{-1}$ of type $i$ are changing the price. Thus, the renewal distribution over types is given by $r_i = \mathcal{T}_i^{-1}\mathbb{E}[\mathcal{T}_i^{-1}]^{-1}$. Let us define $\mathcal{T} = \int_i \mathcal{T}_i r_i di = \mathbb{E}[\mathcal{T}_i^{-1}]^{-1}$.

2. *Calculation of the Aggregate Kurtosis of a Price Change:* To compute the kurtosis notice that the second and the fourth moment of the price change distribution is given by

$$\mathbb{E}[\Delta p^2] = \int_i r_i\mathbb{E}_i[\Delta p^2|\hat{\mu} = 0]di = \mathbb{E}[\mathcal{T}_i^{-1}]^{-1}\int_i \frac{\mathbb{E}_i[\Delta p^2]}{\mathbb{E}_i[\tau]}di = \mathbb{E}[\mathcal{T}_i^{-1}]^{-1}\int_i \sigma_i^2 di = \mathcal{T}\mathbb{E}[\sigma_i^2]$$

$$\mathbb{E}[\Delta p^4] = \int_i r_i\mathbb{E}_i[\Delta p^4|\hat{\mu} = 0]di = \mathbb{E}[\mathcal{T}_i^{-1}]^{-1}\int_i \frac{\mathbb{E}_i[\Delta p^4]}{\mathbb{E}_i[\tau]}di = \mathbb{E}[\mathcal{T}_i^{-1}]^{-1}\int_i \mathcal{K}_i\sigma_i^4\mathcal{T}_i di = \mathcal{T}\mathbb{E}[\mathcal{K}_i\mathcal{T}_i\sigma_i^4]$$

   Therefore the aggregate kurtosis is given by

$$\mathcal{K} = \frac{\mathbb{E}[\Delta p^4]}{\mathbb{E}[\Delta p^2]^2} = \frac{\mathbb{E}[\mathcal{K}_i\mathcal{T}_i\sigma_i^4]}{\mathbb{E}[\sigma_i^2]^2\mathcal{T}} \tag{H.74}$$

3. *Calculation of the Effect of a Monetary Shock:* The effect of a monetary shock is given:

$$\mathcal{M}(\delta) = \int_0^\infty \int_0^1 \hat{y}_{ti} di dt = \int_0^1 \int_0^\infty \hat{y}_{ti} dt di = \int_0^1 \mathcal{M}_i(\delta) di \approx \frac{\mathbb{E}[\mathcal{K}_i\mathcal{T}_i]}{6}\delta \tag{H.75}$$

   Where the last equality comes from Álvarez, Le Bihan and Lippi (2016). Importantly, the effect of a monetary shocks depends on average statistics (computed with the ergodic distribution across firms) and not the aggregate statistics (computed in the previous point with the renewal distribution, which are the ones measured in the data).

4. *Relationship between the Effect of a Monetary Shock and Aggregate Price Statistics:* Using the previous results we have that

$$\mathcal{K}\mathcal{T} = \frac{\mathbb{E}[\mathcal{K}_i\sigma_i^4\mathcal{T}_i]}{\mathbb{E}[\sigma_i^2]^2} = \frac{\mathbb{E}[\sigma_i^4]\mathbb{E}[\mathcal{K}_i\mathcal{T}_i] + \mathbb{C}ov[\mathcal{K}_i\mathcal{T}_i, \sigma_i^4]}{\mathbb{E}[\sigma_i^2]^2} = \frac{\mathbb{E}[\sigma_i^4]\mathcal{M}(\delta)\frac{6}{\delta} + \mathbb{C}ov[\mathcal{K}_i\mathcal{T}_i, \sigma_i^4]}{\mathbb{E}[\sigma_i^2]^2} \tag{H.76}$$

   Solving for $\mathcal{M}(\delta)$, we have the result:

$$\frac{\mathcal{M}(\delta)}{\delta} = \frac{\mathbb{E}_i[\sigma_i^2]^2\mathcal{K}\mathcal{T} - \mathbb{C}ov_i(\mathcal{K}_i\mathcal{T}_i, \sigma_i^4)}{6\mathbb{E}_i[\sigma_i^4]} \tag{H.77}$$

■ The first amplification effect arises from the Jensen inequality applied to the quadratic term $\mathbb{E}_i[\sigma_i^2]^2$. For a given frequency, an increase in the dispersion of uncertainty increases the effect of a monetary shock.

The second effect arises from the covariance term. Firms with low uncertainty behave more like Calvo (low selection effects, high kurtosis) since they are more likely to change the price following a Poisson shock; while firms with high uncertainty behave more like Golosov-Lucas (large selection effect, low kurtosis) since they are more likely to change the price due to the Brownian motion. This relationship between uncertainty and selection effects creates a negative covariance term, which further amplifies monetary effects. Our model with uncertainty cycles can generate these two forces endogenously.

# I Numerical Solution of the Model in Discrete Time

In this section we describe the steps to solve numerically, estimate and compute impulse-response in our model. In this section, we extend our model to include random adjustment probabilities and at the end of the section we analyze the introduction of this property.

## I.1 Model with Perfect Information

The firms' recursive problem with perfect information is given by

$$\begin{aligned}
V(\mu_-) &= \mathbb{E}\left[(1-h)\max_{c,nc}\left\{V^c-\bar{\theta}, V^{nc}(\mu)\right\}+hV^c\right] \\
V^c &= \max_x -x^2+\beta V(x) \\
V^{nc}(\mu) &= -\mu^2+\beta V^{nc}(\mu) \\
\mu &= \mu_- + \sigma_f\epsilon^f+\sigma_u\epsilon^u J \\
\epsilon^u,\epsilon^f &\sim \mathcal{N}(0,1) \qquad J=\begin{cases}1 & \text{with prob. } p\\ 0 & \text{with prob. } 1-p\end{cases}
\end{aligned}$$

If $p=0$ we have a similar model to Golosov and Lucas (2007); and in the case $\sigma_f=0$ we have have a similar model to Gertler and Leahy (2008) and Midrigan (2011). Since we only study these two cases, this problem is equivalent to

$$\begin{aligned}
V(\mu_-) &= \mathbb{E}\left[(1-h)\max_{c,nc}\left\{V^c-\bar{\theta}, V^{nc}(\mu)\right\}+hV^c\right] \\
V^c &= \max_x -x^2+\beta V(x) \\
V^{nc}(\mu) &= -\mu^2+\beta V^{nc}(\mu) \\
\mu &= \mu_- + \sigma\epsilon J \\
\epsilon &\sim \mathcal{N}(0,1) \qquad J=\begin{cases}1 & \text{with prob. } p\\ 0 & \text{with prob. } 1-p\end{cases}
\end{aligned}$$

To solve this recursive problem numerically, we solve the following equivalent problem

$$\begin{aligned}
v(\mu) &= (1-h)\max_{c,nc}\left\{v^c-\bar{\theta}, \mu^2+\beta Ev(\mu)\right\}+hv^c \\
v^c &= \max_x -x^2+\beta Ev(x) \\
Ev(\mu) &= \mathbb{E}\left[v(\mu+\sigma\epsilon J)\right]
\end{aligned}$$

To solve this problem we use function interpolation using 3rd order splines, golden search for the optimization and normal quadrature to compute the expectation. We solve the functional equation with a combination of iterative and colocation methods.

### I.1.1 Model with Perfection Information: Steady-State Price Statistics

We discretize the state space. Given the original distribution of markups in a the compact set $[-\overline{\mu},\overline{\mu}]$, we define a finite set of grid point $\mathcal{M}=\{\mu_1,\mu_2,\ldots,\mu_n\}$. We compute the transition probability $F_{\mathcal{M}'|\mathcal{M}}$ and the ergodic distribution under $F_{\mathcal{M}'|\mathcal{M}}$. Let $n(\mu)$ the ergodic distribution of markups. With the ergodic distribution of markups, we compute the frequency of price change and the distribution of price change using the following formulas

$$N_{ss}=\sum_i \Lambda(\mu_i)n(\mu_i) \quad ; \quad n(\Delta p_j)=\frac{\tilde{n}(\Delta p_j)}{N_{ss}} \quad \tilde{n}(\Delta p_j)=\sum_i \Psi(\Delta p_j|\mu_i)n(\mu_i) \tag{I.78}$$

where $\Lambda(\mu_0)$ is the probability of a price change between periods given the initial state $\mu_i$ and $\Psi(\Delta p_j|\mu_i)$ is the probability of price change $\Delta p_j$ departing from the state $\mu_i$. Notice that by construction $\sum_j \Psi(\Delta p_j|\mu_i) = \Lambda(\mu_i)$. Finally, let $\tilde{F}_{\mathcal{M}'|\mathcal{M}}$ be the transition probability of only no price changes (thus, the sum of columns is less than one). Departing from an initial condition $\tilde{S}_0(0) = 1$, we can compute the survival distribution after t periods of price change recursively

$$\tilde{S}_t(\mu) = \tilde{F}_{\mathcal{M}'|\mathcal{M}} \tilde{S}_{t-1}(\mu) \tag{I.79}$$

Given the survival rate $S_t = \sum_i \tilde{S}_t(\mu_i)$, it is easy to compute the hazard rate.

### I.1.2 Model with Perfection Information: Impulse-Response

Given the initial steady state distribution of markups, $n(\mu)$, we compute the initial distribution after a money shock and before re-pricing $\tilde{n}_0(\mu_i)$

$$\tilde{n}_0(\mu_i) = n(\mu_i + \delta) \tag{I.80}$$

The distribution of markups after re-pricing is given by

$$n_0(\mu_i) = \tilde{n}(\mu_i) I(\mu_i \in [-\overline{\mu}, \overline{\mu}]) + I(\mu_i = 0) \left( \sum_i \tilde{n}(\mu_i) I(\mu_i \notin [-\overline{\mu}, \overline{\mu}]) \right) \tag{I.81}$$

With the initial distribution we compute the distribution of markups at period t, $n_t(\mu)$, with following stochastic process

$$\begin{aligned} \mu_{1+1} &= \begin{cases} 0 & \text{if } J^h_{t+1} = 0 \textit{ and } v^c - \theta \geq v^{nc}(\bar{\mu}_{t+1}) \textit{ or } J^h_{t+1} = 1 \\ \bar{\mu}_{t+1} & \text{otherwise} \end{cases} \\ \bar{\mu}_{1+1} &= \mu_t + \sigma \epsilon_{t+1} J^q_{t+1} \end{aligned}$$

where $J^p_{t+1} = 1$ with probability $p$ and $J^h_{t+1} = 1$ with probability $h$. Both variables are i.i.d. across time. Once we have the distribution of markups at period $t$, we can compute output gap and frequency of price change

$$Y_t = -\sum_i \mu_i n_t(\mu_i) \quad ; \quad N_t = \sum_i \Lambda(\mu_i) n_t(\mu_i) \tag{I.82}$$

## I.2 Model with Imperfect Information

The firms' recursive problem with imperfect information is given by

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

We skip the derivation of this recursive problem (see section G to derive it). To solve this recursive problem numerically, we solve the following equivalent problem

$$
\begin{aligned}
v(\hat{\mu}, \Sigma) &= (1-h)\max_{c,nc}\left\{v^c(\Sigma) - \bar{\theta}, \hat{\mu}^2 + \beta E v(\hat{\mu}, \Sigma)\right\} + h v^c(\Sigma) \\
v^c(\Sigma) &= \max_{x} -x^2 + \beta E v(x, \Sigma) \\
Ev(\hat{\mu}_-, \Sigma_-) &= \mathbb{E}\left[v(\hat{\mu}_- + \frac{\Sigma}{\sqrt{\Sigma+\gamma^2}}\epsilon, \frac{\gamma^2\Sigma_-}{\gamma^2+\Sigma_-} + \sigma_f^2 + \sigma_u^2 J)\right]
\end{aligned}
$$

To solve this problem we use function interpolation using 3rd order splines, golden search for the optimization and normal quadrature to compute the expectation.

To solve this problem we use function interpolation using 3rd order splines, golden search for the optimization and normal quadrature to compute the expectation. We solve the functional equation with a combination of iterative and colocation methods.

#### I.2.1 Model with Imperfection Information: Steady-State Price Statistics

We use a discretization of the state space. Given the original distribution of markups in a the compact set

$$
[-\max_{\Sigma\in[\underline{\Sigma},\overline{\Sigma}]}\overline{\mu}(\Sigma), \max_{\Sigma\in[\underline{\Sigma},\overline{\Sigma}]}\overline{\mu}(\Sigma)] \times [\underline{\Sigma}, \overline{\Sigma}] \tag{I.83}
$$

we define a finite set of grid point $\mathcal{X} = \{\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_n\} \times \{\Sigma_1, \Sigma_2, \ldots, \Sigma_m\}$. We compute the transition probability $F_{\mathcal{X}'|\mathcal{X}}$ and the ergodic distribution under $F_{\mathcal{X}'|\mathcal{X}}$. Let $n(\hat{\mu}, \Sigma)$ the ergodic distribution of markups and uncertainty. With the ergodic distribution of markups and uncertainty, we compute the frequency of price change and the distribution of price change using the following formulas

$$
N_{ss} = \sum_i \Lambda(\hat{\mu}_i, \Sigma_i) n(\hat{\mu}_i \Sigma_i) \quad ; \quad n(\Delta p_j) = \frac{\tilde{n}(\Delta p_j)}{N_{ss}} \quad \tilde{n}(\Delta p_j) = \sum_i \Psi(\Delta p_j | \hat{\mu}_i \Sigma_i) n(\hat{\mu}_i \Sigma_i) \tag{I.84}
$$

where $\Lambda(\hat{\mu}_0, \Sigma_0)$ is the probability of a price change between periods given the initial state $(\hat{\mu}_i, \Sigma_i)$ and $\Psi(\Delta p_j|\hat{\mu}_i, \Sigma_i)$ is the probability of price change $\Delta p_j$ departing from the state $(\hat{\mu}_i, \Sigma_i)$. Finally, let $\tilde{F}_{\mathcal{X}'|\mathcal{X}}$ be the transition probability of only no price changes (thus, the sum of columns is less than one). Departing from the renewal distribution $\tilde{S}_{t-1}(0, \Sigma) = R(\Sigma)$, we can compute the survival distribution after $t$ periods of price change recursively

$$
\tilde{S}_t(\hat{\mu}) = \tilde{F}_{\mathcal{X}'|\mathcal{X}} \tilde{S}_{t-1}(\hat{\mu}) \tag{I.85}
$$

Given the survival rate $S_t = \sum_i \tilde{S}_t(\hat{\mu}_i)$, it is easy to compute the hazard rate.

#### I.2.2 Model with Imperfection Information: Impulse-Response

Given the steady state distribution of markups gap estimate and uncertainty, $n(\hat{\mu}, \Sigma)$, we compute the initial distribution of markups gap estimate, forecast error and uncertainty after a money shock. This distribution depends of three parameters: the size of the monetary shock $\delta$; the size of the uncertainty shock of size $\kappa\overline{\Sigma}$; and the observe fraction of the money shock incorporate in the estimate $\alpha$. Given these three parameters the distribution of distribution of markups gap estimate, uncertainty and forecast error after a money shock and before re-pricing $\tilde{n}_0(\mu_i, \Sigma_i, \varphi_i)$ is given by

$$
\tilde{n}_0(\mu_i, \Sigma_i, \varphi_i) = n(\mu_i + \alpha\delta, \Sigma_i - \kappa\overline{\Sigma}) I\left(\varphi_i = -(1-\alpha)\delta\right)
$$

The distribution after re-pricing is given by

$$n_0(\mu_i, \Sigma_i, \varphi_i) \quad = \quad \tilde{n}_0(\mu_i, \Sigma_i, \varphi_i) I(\mu_i \in [-\overline{\mu}(\Sigma_i), \overline{\mu}(\Sigma_i)]) + I(\mu_i = 0) \left( \sum_i \tilde{n}_0(\mu_i, \Sigma_i, \varphi_i) I(\mu_i \notin [-\overline{\mu}(\Sigma_i), \overline{\mu}(\Sigma_i)]) \right)$$

Given the initial state the stochastic process for the firm state is given by

$$\begin{aligned}
\mu_{1+1} &= \begin{cases} 0 & \text{if } J^h_{t+1} = 0 \textit{ and } v^c(\Sigma_{t+1}) - \theta \geq v^{nc}(\bar{\mu}_{t+1}, \Sigma_{t+1}) \textit{ or } J^h_{t+1} = 1 \\ \bar{\mu}_{t+1} & \text{otherwise} \end{cases} \\
\bar{\mu}_{1+1} &= \mu_t + \frac{\Sigma_t}{\sqrt{\Sigma_t + \gamma^2}} \epsilon_{t+1} + \varphi_t \frac{\Sigma_t}{\Sigma_t + \gamma^2} \\
\Sigma_{1+1} &= \frac{\gamma^2}{\Sigma_t + \gamma^2} \Sigma_t + \sigma_f^2 + \sigma_u^2 J^p_{t+1} \\
\varphi_{1+1} &= \varphi_t \frac{\gamma^2}{\Sigma_t + \gamma^2}
\end{aligned}$$

Once we have the distribution of firms at period $t$, we can compute forecast error, output gap and frequency of price change

$$\begin{aligned}
\mathcal{F}_t &= \sum_i \varphi_i n_t(\mu_i, \Sigma_i, \varphi_i) \\
\mathcal{Y}_t &= -\sum_i (\hat{\mu}_i + \varphi_i) n_t(\mu_i, \Sigma_i, \varphi_i) \\
\mathcal{N}_t &= \sum_i \Lambda(\mu_i, \Sigma_i) n_t(\mu_i, \Sigma_i, \varphi_i)
\end{aligned}$$

# J Robustness of the Dynamic Effect of a Monetary Shock

In the main text, we calibrate the model using the evidence in Zbaracki *et al.* (2004) and Levy *et al.* (1997) for the size of the menu cost; and we use hazard rate in Nakamura and Steinsson (2008). Since there is a reasonable amount of empirical disagreement about these two target, this section presents robustness exercises with these two targets. Subsection J.1 presents robustness with respect to the size of the menu cost and J.2 presents robustness with respect the hazard rate.

## J.1 Robustness with Respect to the Menu Cost

This subsection re-calibrates the model for different menu cost and it shows the changes in the real effect of a monetary shock under the these calibration. Figure XIX shows when we calibrate the menu cost 20% higher or lower than the baseline case in the main text and table VIII shows the price statistics. Notice

Table VIII: Parameters and Moments For Different Menu Cost

| | (1) Baseline | (2) Low Menu Cost | (3) High Menu Cost |
|---|---|---|---|
| **Parameters** | | | |
| $\theta$ | 0.064 | 0.051 | 0.077 |
| $\sigma_f$ | 0 | 0 | 0 |
| $\sigma_u$ | 0.198 | 0.198 | 0.198 |
| $\lambda$ | 0.016 | 0.016 | 0.016 |
| $\gamma$ | 0.233 | 0.233 | 0.233 |
| **Moments** | | | |
| $\mathbb{E}[\tau]$ in months | 10 | 10 | 9.9 |
| std$[\lvert\Delta p\rvert]$ | 0.07 | 0.07 | 0.07 |
| hazard rate slope | $-0.009$ | $-0.009$ | $-0.009$ |
| kurtosis$[\Delta p]$ | 1.98 | 1.96 | 1.94 |

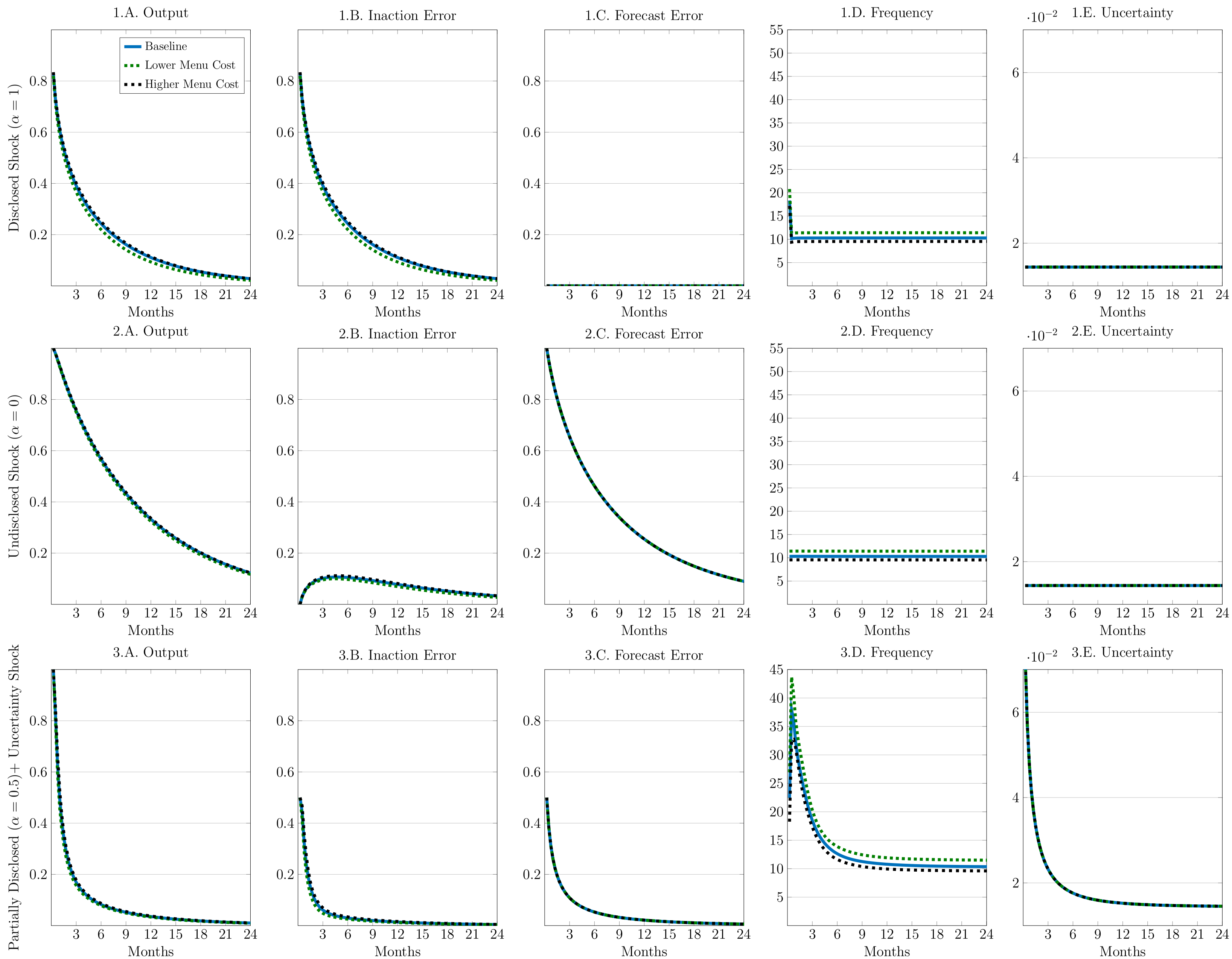
Models: (1) Baseline: as the main text; (2) Menu cost 20% lower than baseline; (3) Menu cost 20% higher than the baseline.

## J.2 Robustness with Respect to the Hazard Rate

This subsection re-calibrates the model to different hazard rate and it shows the changes in the real effect of a monetary shock under the these calibration. We cover 4 cases: (i) Baseline, where we repeat the calibration in the main text with the target of Nakamura and Steinsson (2008); (ii) Constant, where we repeat the calibration with perfect information and fat-tail shocks; (iii) Increasing-decreasing, where we recalibrate a zero slope of the hazard rate with first increasing and then decreasing; (iv) Highly decreasing, where we duplicate the slope of the hazard rate by from -0.009 to -0.012. Since we cannot generate a decreasing and then increasing cases, we do not cover this case. Table IX shows each calibration.

Figure XX shows the hazard rate in the different cases, together with the renewal and marginal distribution of uncertainty in the fourth cases (notice that with $\gamma = 0$ the distribution of uncertainty is trivially 0 and therefore we don't show it). As we can see in Panel C, the increasing-decreasing case has a lower dispersion uncertainty—in all firm and in all firms conditional of adjustment—than in the benchmark case. The opposite direction is the highly-decreasing hazard rate. In this case, as we can see in panel D, there is a large dispersion of uncertainty across firms and even a highly difference in between the renewal and the marginal distribution of uncertainty. In conclusion, we can see clearly

Figure XIX: Impulse-Response Function for Different Calibrations of the Menu Cost




Note: Impulse-responses after a monetary shock. The first row plots the impulse-response with a fully disclosed monetary shock ($\alpha = 1$), the second row plots the impulse-response with a fully undisclosed shock ($\alpha = 0$), and the third row with a partially disclosed ($\alpha = 0.5$) shock together with a large uncertainty shock.

Table IX: Parameters and Targets

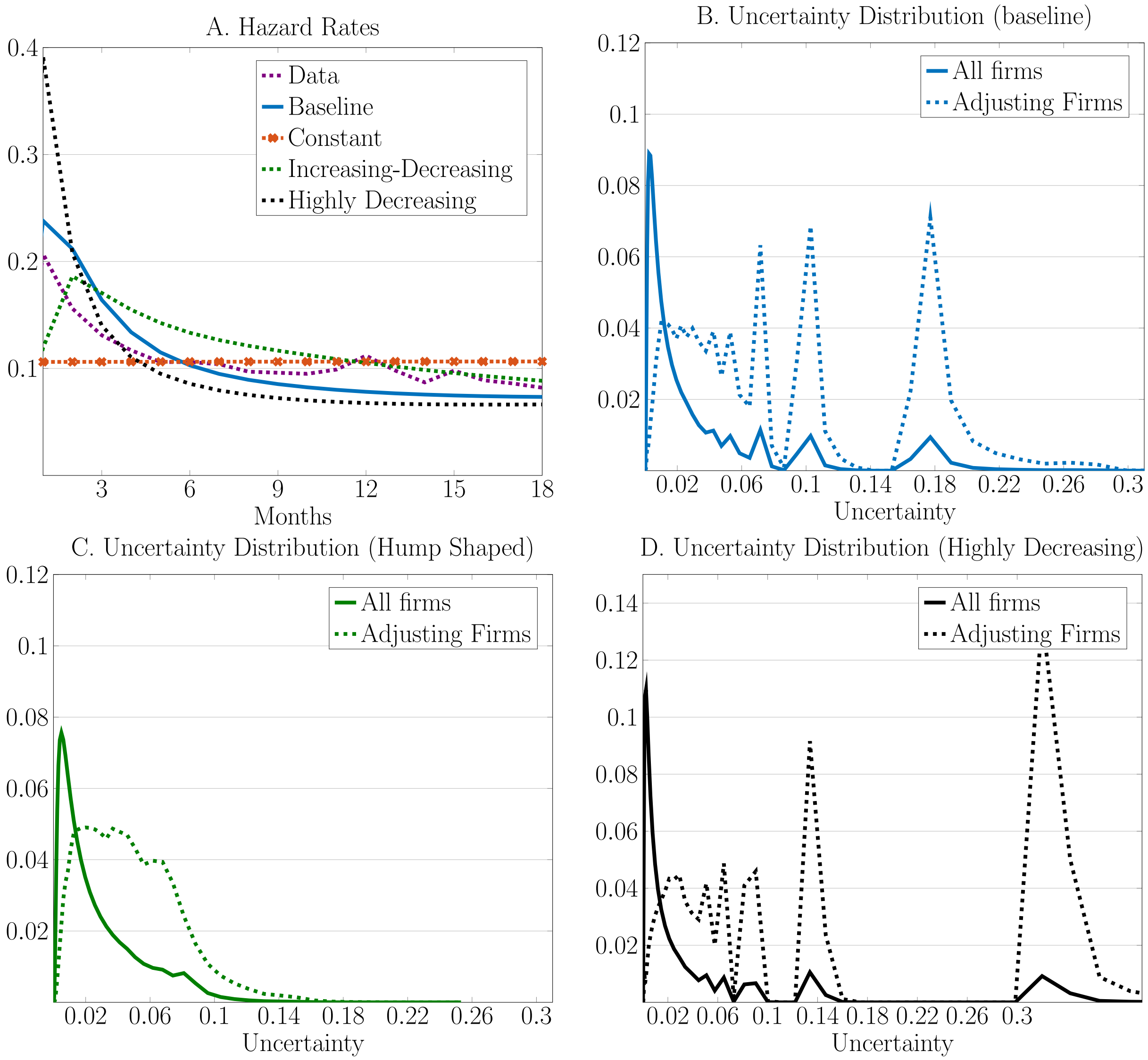
| | (1) Baseline | (2) Constant | (3) Increasing-Decreasing | (4) Highly decreasing |
|---|---|---|---|---|
| **Parameters** | | | | |
| $\sigma_f$ | 0 | 0 | 0 | 0 |
| $\sigma_u$ | 0.198 | 0.146 | 0.198 | 0.2729 |
| $\lambda$ | 0.016 | 0.013 | 0.015 | 0.0134 |
| $\gamma$ | 0.233 | 0 | 0.500 | 0.233 |
| **Moments** | | | | |
| $\mathbb{E}[\tau]$ in months | 10 | 10 | 10 | 10 |
| std[$\lvert\Delta p\rvert$] | 0.07 | 0.08 | 0.05 | 0.10 |
| hazard rate slope | $-0.009$ | 0 | $-0.001$ | $-0.019$ |
| kurtosis[$\Delta p$] | 1.96 | 2.226 | 1.7416 | 2.85 |

Models: (1) Baseline: Same calibration as the main text; (2) Perfect info and fat-tailed shocks; (3) Heterogeneous uncertainty. Hazard rate's slope - average between 1 and 18 months.

how the slope of the hazard rate implies different dispersion in the uncertainty distribution across firms.

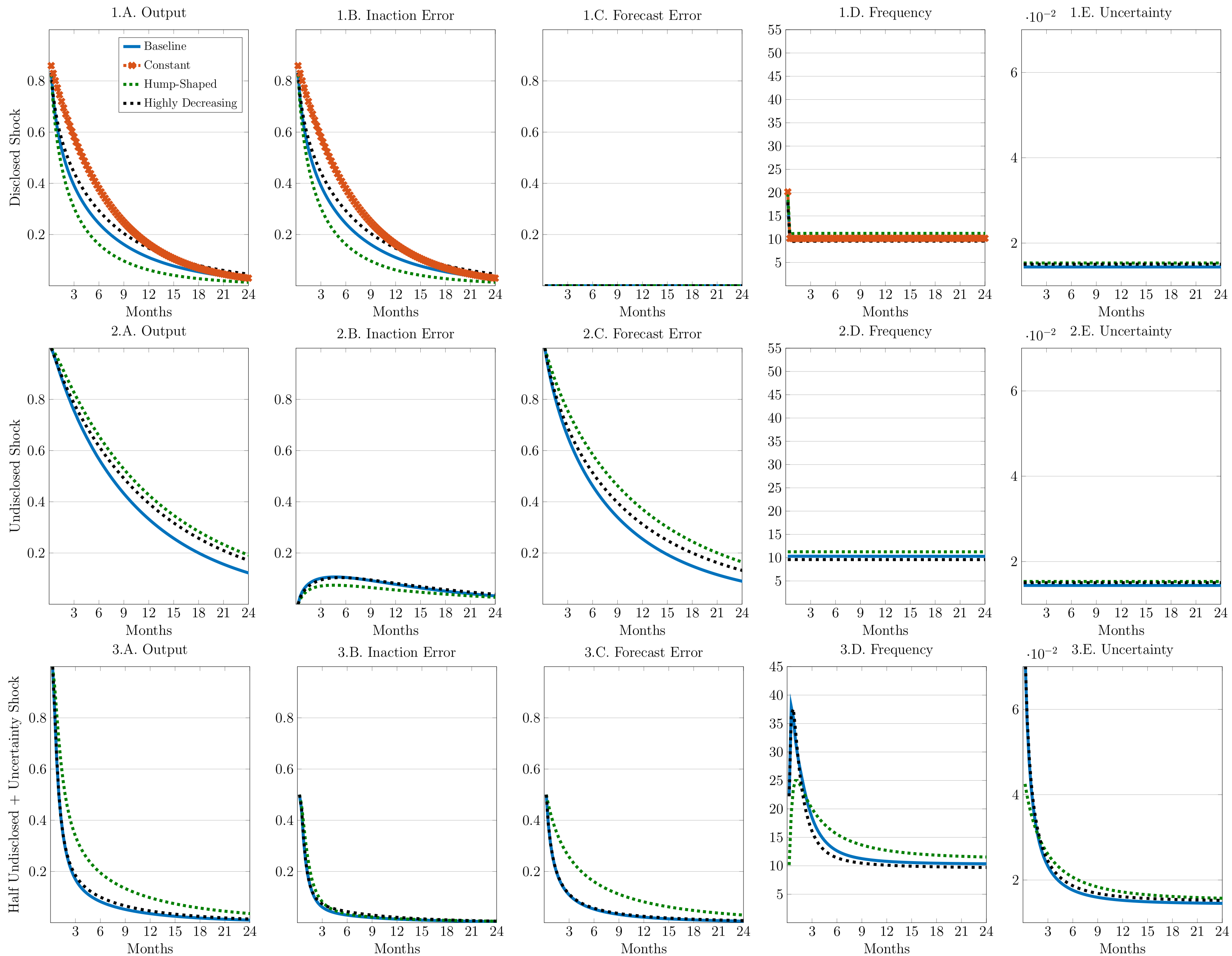
Figure XXI shows the impulse-response functions of output, inaction errors, forecast errors, frequency of price change and average uncertainty dynamics. From these figures we can extract the following conclusion: i) lower heterogeneity in uncertainty implies lower effect of monetary shocks in the disclosed monetary shock, ii) these relation is reverted in the case of undisclosed monetary shocks, since the amplification given by the dispersion of the forecast error is cancel with a higher $\gamma$.

Figure XX: Hazard Rate for Different Values of $\lambda$




Panel A plots the hazard rates for different calibration and int he Data. Panels B to D show the distribution of uncertainty in the baseline case as in the main text, increasing-decreasing hazard rate and in the highly decreasing hazard rate, respectively.

Figure XXI: Impulse-Response Function for Different Calibrations of the Hazard Rate




Note: Impulse-responses after a monetary shock. The first row plots the impulse-response with a disclosed monetary shocks, the second row plots the impulse-response with an undisclosed shocks and the final row with partially disclosed row together with an large uncertainty shock as in the main text.

## K CalvoPlus Model with Imperfect Information

In this section we extend our model to include random menu costs. It is a version of the CalvoPlus model as in Nakamura and Steinsson (2008). With this model, we can fix two shortcomings of our benchmark setup: (1) we are able to obtain small price changes, and (2) we reduce the elasticity of the adjustment frequency to uncertainty. Moreover, in this case our model generates more propagation of nominal shocks than Gertler and Leahy (2008). We target a kurtosis of 4, and the same statistics as the main text. We also target the physical cost of changing the price, taking into account that there are some free price changes. Table X describes the parameters in the CalvoPlus model. As we can see, now all the models can match the kurtosis of price change with a different ratio between free-price changes and costly price changes.

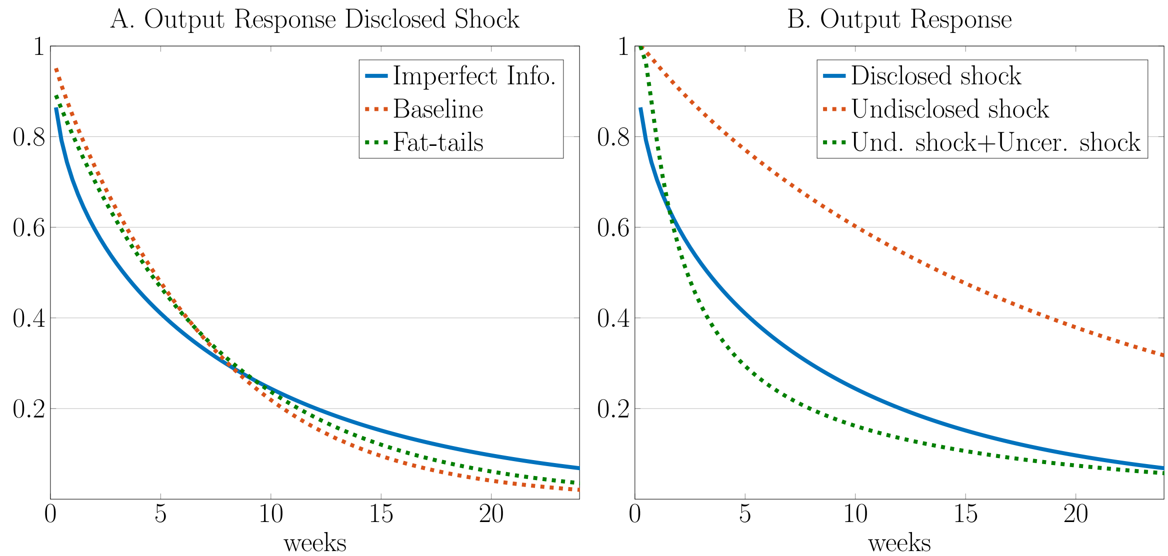
Table X: Model Parameters and Data Targets

| | US Data | (1) Benchmark | (2) Fat-tailed shocks | (3) Imperfect Info |
|---|---|---|---|---|
| **Parameters** | | | | |
| $\theta$ | | 0.06994 | 0.0647 | 0.0647 |
| $\sigma_f$ | | 0.0037 | | 0 |
| $\sigma_u$ | | | 0.1236 | 0.29 |
| $\lambda$ | | | 0.01960 | 0.0085 |
| $\gamma$ | | | | 0.3 |
| $h$ | | 0.0236 | 0.0120 | 0.0065 |
| **Moments** | | | | |
| $\mathbb{E}[\tau]$ in months | 10 | 10* | 10* | 10* |
| std[$\lvert\Delta p\rvert$] | 0.08 | 0.0158 | 0.08* | 0.101* |
| Hazard rate's slope | $-0.007$ | 0.0009 | 0 | $-0.012$* |
| kurtosis[$\Delta p$] | 3.95 | 3.938* | 3.9585* | 3.70* |
| Cost of price change | 0.0064 | 0.0068* | 0.0064* | 0.0060* |

Data: CPI Data from US.
Models: (1) Benchmark: Perfect info with only frequent shocks; (2) Perfect info with only fat-tailed shocks; (3) Imperfect info with fat-tailed shocks. Hazard rate's slope is the average between 1 and 18 months. Asterisk = target moment in calibration.

Figure XXII describes the impulse-response for the same experiments we did in the main text. Panel A describes three cases: Perfect information benchmark, perfect information and fat-tailed shocks and our model with disclosed shocks. The impulse-responses are similar, with the qualitative property that the model with uncertainty cycles generates less real effect on impact but more persistence. The last property is the same as in the main text without CalvoPlus. Panel B describes our model with disclosed monetary shock, undisclosed monetary shock and undisclosed monetary shock with an uncertainty shock. As we can see, when forecast errors are calibrated to match micro-price statistics, there are large effects of monetary shocks that are dampened when uncertainty is high.

Figure XXII: Output Response to a Monetary Shock for CalvoPlus Model




Left panel: output response for a disclosed shock in the three models with random menu cost. Right panel: output response for different shocks in CalvoPLus model with uncertainty cycles.

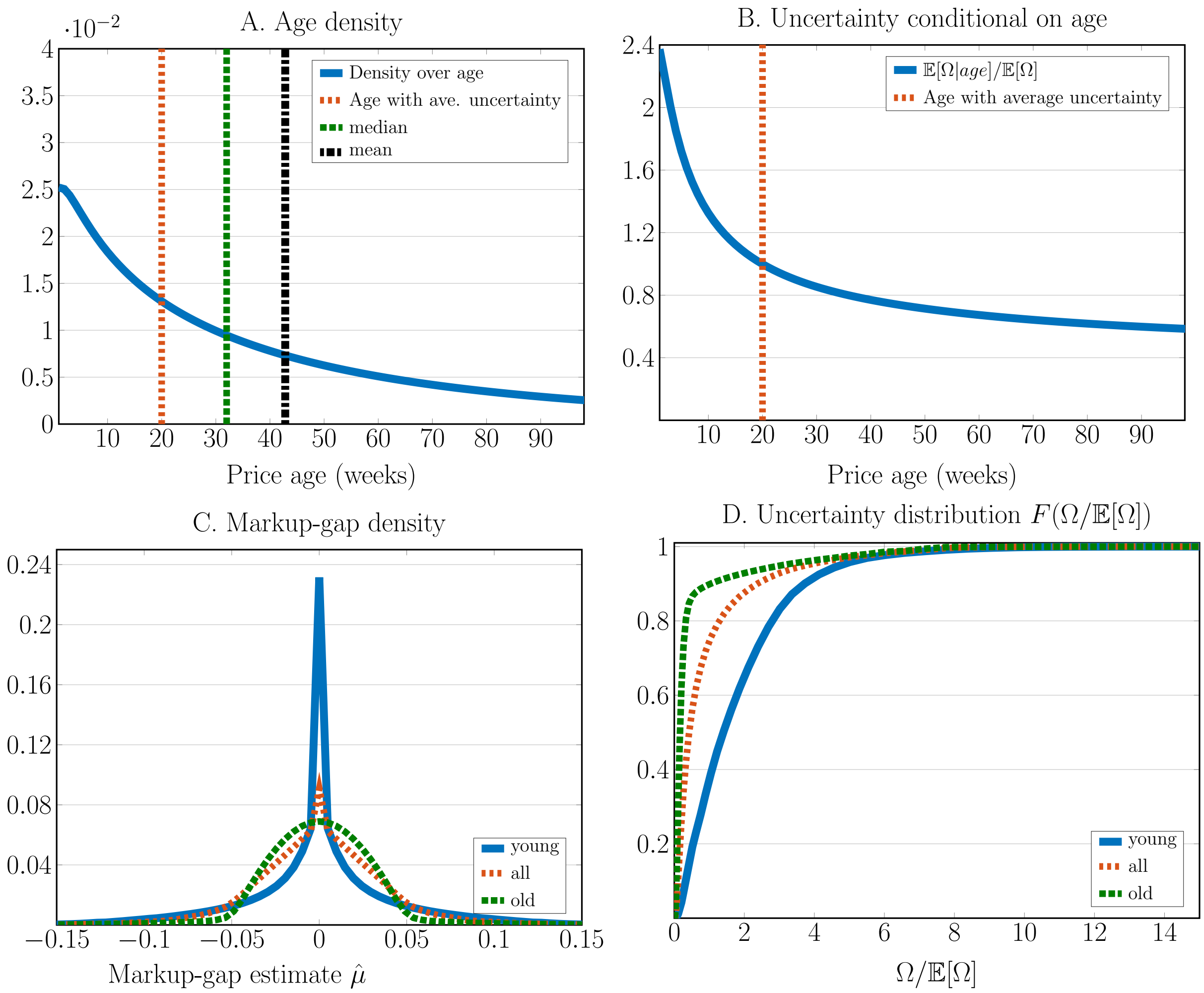
# L Overidentifying Restrictions: Age and Price Statistics

In this section, we conduct robustness checks of our calibration by exploiting the tight connection between the age of price and firm uncertainty produced by our model. Price age is measured as the number of periods that a price has remained unchanged. In particular, our model predicts that young (uncertain) prices are larger, more dispersed, and more likely to be reset than older (certain) prices. These predictions are documented by Campbell and Eden (2014) using weekly scanner data. We use this evidence as an over-identifying restriction for our model.

## L.1 Relationship between uncertainty $\Omega$ and age $a$

Figure XXIII describes the relation age and uncertainty. Panel A plots the ergodic density of price age. Panel B plots the ratio of conditional to unconditional uncertainty $\mathbb{E}[\Omega|a]/\mathbb{E}[\Omega]$. A price with uncertainty equal to average uncertainty ($\mathbb{E}[\Omega] = 0.0026$) has an age of 20 weeks, whereas the median and average age are 32 and 42 respectively. Young prices have almost twice the average level of uncertainty, whereas old prices have about 75% of the average level of uncertainty. Panel C plots the distribution of prices conditional on age. Young prices are more dispersed than old prices. Panel D plots the accumulated distribution of uncertainty normalized by the mean. The distribution of uncertainty of young prices first-order stochastically dominates the distribution of old prices. Young prices are twice as flexible as old prices.

Figure XXIII: Age dependent statistics and uncertainty

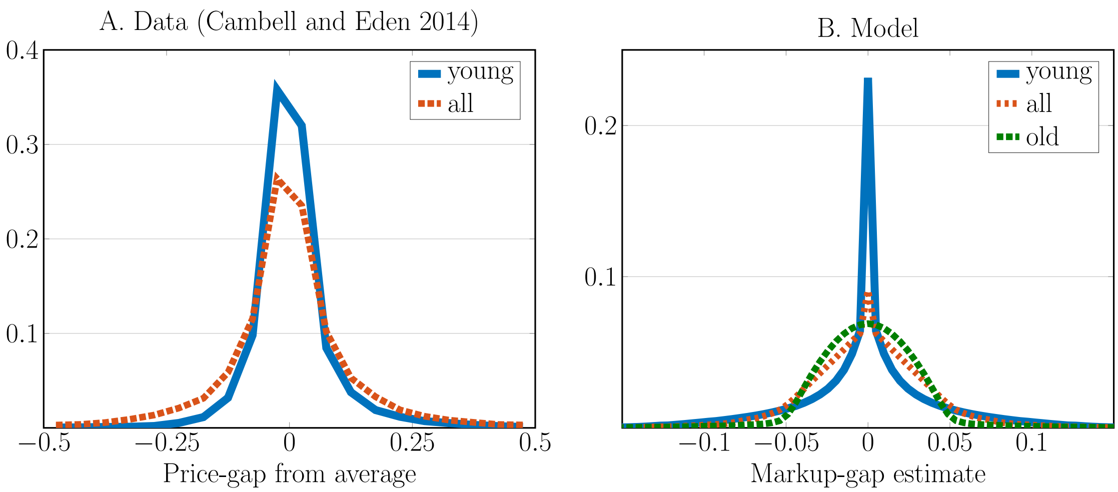


## L.2 Scanner data vs. Model

Campbell and Eden (2014) define a young price if its age is less than three weeks and an old price if its age is more than four weeks. That paper finds that conditional on adjustment, young prices have double the dispersion of old prices (7% vs. 15%) and that price changes in the extreme tails of the price change distribution tend to be young. Regarding the frequency, it finds that young prices are three times more likely to be changed than old prices (36% vs 13%).

We compute analogous numbers in our model, defining young prices to be in the 20th quintile of the price age distribution (2 months) and old prices to be in the 80th quintile (12 months). We obtain that price dispersion is 1.5 times higher for young prices compared to old prices, and that adjustment frequency is 2 times higher for young prices. Interestingly, the uncertainty faced by young prices is also three times the uncertainty faced by old prices, thus the relative adjustment frequency seems to be informative about the relative uncertainty faced by firms. See Figure XXIV.

Figure XXIV: Price Gap Distribution (Data vs. Model)




| Statistic | Scanner data | | | | Model with heterogenous uncertainty | | | |
|---|---|---|---|---|---|---|---|---|
| | All | Young | Old | Ratio | All | Young | Old | Ratio |
| Frequency % | 15 | 36 | 13 | 2.8 | 10.00 | 13.12 | 6.44 | 2.0 |
| std(Price gap) | 11.4 | 15.1 | 6.9 | 2.2 | 3.61 | 4.0 | 3.0 | 1.33 |
| Uncertainty*100 | | | | | 0.014 | 0.027 | 0.008 | 3.3704 |

## L.3 Summary of price statistics by age

Table XI: Micro-Price Statistics in Calibrated Model

| moments | All | Young | Old |
|---|---|---|---|
| Absolute Value of Price Change | | | |
| Mean | 0.130 | 0.157 | 0.1258 |
| Std Deviation | 0.059 | 0.045 | 0.069 |
| Skewness | 0.905 | 1.493 | 0.580 |
| Kurtosis | 4.266 | 7.448 | 2.910 |
| Price Change | | | |
| Mean | -0.00 | -0.00 | -0.00 |
| Std Deviation | 0.029 | 0.164 | 0.143 |
| Skewness | 0.039 | -0.012 | 0.029 |
| Kurtosis | 1.891 | 1.449 | 2.060 |

This table describes micro-price statistics for all, young and old prices.

# M Impact Effects: Continuous vs. Leptokurtic Processes

This section compute the impulse-response after a monetary shock of the Golosov and Lucas (2007) and the Gertler and Leahy (2008) models up to first order. We show that: i) the impact effect on output of a monetary shock in the leptokurtic model of Gertler and Leahy (2008) is *lower* than the Brownian model in Golosov and Lucas (2007), ii) the persistence of the output effects generated in Gertler and Leahy (2008) is higher than in Golosov and Lucas (2007) for a sufficiently small menu cost.

The environment is as in the main text. The markup gap law of motion is given by

$$d\mu_t = \sigma_f dW_t + z_t dQ_t, \tag{M.86}$$

here $W_t$ is a Brownian motion, $z_t$ is an uniform distribution with support $[-\frac{\phi}{2}, \frac{\phi}{2}]$ and $Q_t$ is a Poisson counter with intensity $\lambda$. For simplicity we assume a positive money shock $\delta$.

**Impulse-response function in Gertler and Leahy 2008:** We proceed to compute the impulse-repulse response in the case with $\sigma_f = 0$. For simplicity we assume that $\frac{\phi}{2} > 2\sqrt{\theta(r+\lambda)}$. The HJB is given by

$$(r+\lambda)V(\mu) = -\mu^2 + \lambda \mathbb{E}_\phi \left[\max\left\{V(\mu+\phi), V(0)-\theta\right\}\right]. \tag{M.87}$$

Given that $\phi$ is uniform, the $\mathbb{E}_\phi \left[\max\left\{V(\mu+\phi), V(0)-\theta\right\}\right]$ is independent of $\mu$. With this property, it is easy to show that the continuation region is given by $[-\overline{\mu}_u, \overline{\mu}_u]$ with $\overline{\mu}_u = \sqrt{\theta(r+\lambda)}$.

The steady state distribution of markups is given by an uniform distribution with a probability atom at $\mu = 0$. The probability atom $P$ in steady state satisfies the zero net entry. The exit rate is given by $P\lambda dt$, where $P$ is the mass as zero and $\lambda dt$ is the probability of exit in a time period $dt$. The entry rate is given by $\lambda dt[1 - \frac{2\overline{\mu}}{\phi}]$, where 1 is the measure of all firms and $[1-\frac{2\overline{\mu}}{\phi}]$ is the probability of a price change conditional of receiving an idiosyncratic shock. Given that the entry rate has to be equal to the exit rate, we have that $P = \left[1 - \frac{2\overline{\mu}}{\phi}\right]$. Since there is a measure 1 of firms, the distribution of markups for all $\mu \notin [-\overline{\mu}_u, \overline{\mu}_u]$ is given by

$$f(\mu) = \frac{1}{\phi}, \tag{M.88}$$

with a probability atom $P = \left[1 - \frac{\overline{\mu}}{\phi}\right]$ in $\mu = 0$.

To compute the impulse-response we proceed in two steps: 1) the impact effect, and 2) the dynamic effect. The impact effect is given by the decrease of markup of size $\delta$ minus the average price change $-\int_{-\overline{\mu}-\delta}^{-\overline{\mu}} x\frac{1}{\phi}dx = \delta\frac{1}{\phi}[\overline{\mu} - \delta/2]$, therefore the impact effect is given by $\delta\left[1 - \frac{1}{\phi}\left(\overline{\mu} + \delta/2\right)\right]$ and $\hat{Y}_0 = \delta\left[1 - \frac{1}{\phi}\left(\overline{\mu} + \delta/2\right)\right]$. Since conditional of receiving an idiosyncratic shock, the new markup has average zero, the mean reversion of the output is given by $\lambda dt$, with the impulse response given up to a first order

$$\frac{\hat{Y}_t}{\delta} = e^{-\lambda t}\left[1 - \frac{\overline{\mu} + \delta/2}{\phi}\right] = e^{-\lambda t}\left[1 - \frac{\overline{\mu}}{\phi}\right] + o(\delta) \tag{M.89}$$

Notice that the impact effect is linear in $\delta$ and the persistence is given by $\lambda$.

**Impulse-response function in Golosov and Lucas 2007:** For computing the $\lambda = 0$, we follow

Kolkiewicz (2002) and Álvarez and Lippi (2014)'s Online Appendix for the case $n = 1$. In this case

$$\frac{\hat{Y}_t}{\delta} = 1 - \frac{\mu_f}{\delta} \sum_{k=1}^{\infty} \frac{1 - \exp\left(-k^2\pi^2 t/(2\mathbb{E}[\tau])\right)}{k\pi} \times \dots \tag{M.90}$$

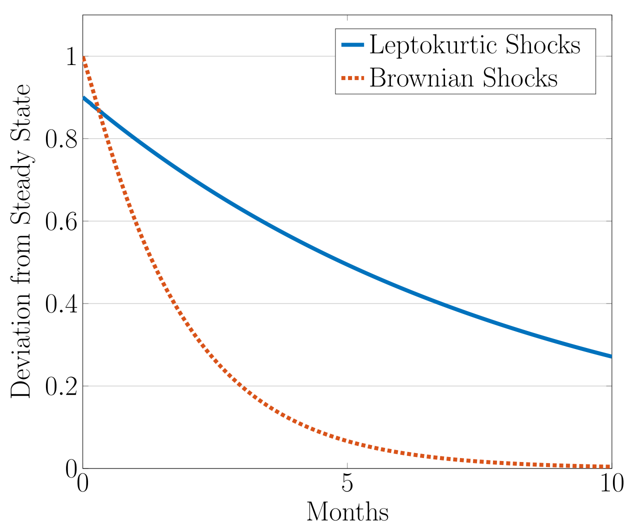
$$\dots + \int_{-1}^{-\delta/\mu_f} \left[\sin(k\pi(1+x)) - \sin(k\pi(1-x))\right] \times \lambda(x, \frac{\delta}{\mu_f}, 1)dx + o(\delta) \tag{M.91}$$

$$\lambda(x,\delta,\mu_f) = \begin{cases} \frac{1}{\mu_f}\left(1 + \frac{\delta}{\mu_f} + \frac{x}{\mu_f}\right) & \text{if } \frac{x}{\mu_f} \in \left[-1 - \frac{\delta}{\mu_f}, -\frac{\delta}{\mu_f}\right] \\ \frac{1}{\mu_f}\left(1 - \frac{\delta}{\mu_f} - \frac{x}{\mu_f}\right) & \text{if } \frac{x}{\mu_f} \in \left(-\frac{\delta}{\mu_f}, 1 - \frac{\delta}{\mu_f}\right] \end{cases} \tag{M.92}$$

$$\mu_f = \left(6\theta\sigma_f^2\right)^{1/4} \tag{M.93}$$

**Discussion:** Figure XXV shows that the price response is larger with leptokurtic shocks, which implies that the output effects are smaller (we follow the same calibration as the main text with $\phi = 0.25$ as in Gertler and Leahy (2008)). We observe that the impact effect on output is lower for leptokurtic shocks, but the persistence is lower for the Brownian case, because the infinite sum forces the exponential function to converge faster to zero.

Figure XXV: Output Impulse-Response with Leptokurtic and Brownian Shocks



# N Intensive and Extensive Margin Decomposition

In this section we analyze a decomposition of the output effects into an intensive and extensive margin. The strategy is as follows:

1.1 We revisit Caballero and Engel (2007)'s standard decomposition of price adjustments after a monetary shock into extensive and intensive margins; we show that the nature of their decomposition—static and to a first order—is problematic for a continuous time framework as ours.[5] We exemplify this through a simple menu cost model with Brownian motion and then in our model with information frictions.

1.2 As an alternative, we propose a decomposition in Caballero and Engel (2007)'s spirit that computes the extensive margin as a residual. We explain how to implement this decomposition for the impact response and then a dynamic response. Since it does not rely on first order effects, it is meaningful for our model.

1.3 We implement the dynamic decomposition for the case of disclosed (fully observable) nominal shocks.

1.4 We implement the dynamic decomposition for the case of partially disclosed nominal shocks and discuss some challenges.

1.5 We quantify the relative importance of the two margins for the same set of exercises in the main text.

## N.1 Challenges with Standard Decomposition

We follow closely the notation and strategy in Caballero and Engel (2007). Consider an economy with a continuum of firms represented by an idiosyncratic state $\mu$ (markup gaps) and let $f(\mu)$ be its steady state cross-sectional density. Let $\Lambda(\mu)$ be the probability of price adjustment given the state $\mu$, or the adjustment hazard, which is increasing. With these objects at hand, we can express the aggregate price adjustment when there is a small monetary shock of size $\Delta m = \delta > 0$, denoted by $\Delta p(\delta)$, as follows:[6]

$$\Delta p(\delta) = -\int (\mu - \delta)\Lambda(\mu - \delta) f(\mu)\, d\mu. \tag{N.94}$$

After a first order Taylor approximation around $\delta = 0$ we obtain

$$\frac{\Delta p(\delta)}{\delta} = \underbrace{\int \Lambda(\mu) f(\mu)\, d\mu}_{Intensive\ \mathcal{I}} + \underbrace{\int \mu \Lambda'(\mu) f(\mu)\, d\mu}_{Extensive\ \mathcal{E}} + o(\delta) \tag{N.95}$$

The first term $\mathcal{I}$, called the *intensive margin*, corresponds to the adjustment frequency and describes the additional price increase resulting from a *change in the size* of the adjustment in firms that were going to adjust anyway. In contrast, the second term $\mathcal{E}$, called the *extensive margin*, reflects the price changes resulting from a *change in the fraction* of adjusters. Note that these quantities reflect *impact* effects and are static measures.

[5] In their discussion about the extensive margin, Caballero and Engel (2007) recognizes that their extensions with idiosyncratic shocks "[...] are more easily implemented in discrete time."

[6] The negative sign comes from the fact that a positive markup gap implies a negative desired price change.

**Application to Brownian state** Before applying this decomposition into our model, we consider a simple menu cost model where the state follows a Brownian motion and the inaction region is the interval $\mathcal{R} \equiv [\underline{\mu}, \overline{\mu}]$. The adjustment hazard is given by $\Lambda(\mu) = \Delta(\mu - \underline{\mu}) + \Delta(\mu - \overline{\mu})$, where $\Delta$ denotes the Dirac delta function around zero.[7] It is easy to see that the probability mass at either boundary of the inaction region $\partial\mathcal{R} = \{\underline{\mu}, \overline{\mu}\}$ is equal to zero, because as soon as the state touches the boundary the price is reset and mass cannot accumulate at those points. Therefore, $f(\underline{\mu}) = f(\overline{\mu}) = 0$. Putting things together, and using properties of the Dirac delta[8], we obtain

$$\begin{aligned}\frac{\Delta p(\delta)}{\delta} &= \int \left[\Delta(\mu - \underline{\mu}) + \Delta(\mu - \overline{\mu})\right] f(\mu)\ d\mu + \int \mu\Lambda'(\mu) f(\mu)\ d\mu + o(\delta) \\ &= \int \Delta(\mu - \underline{\mu}) f(\underline{\mu})\ d\mu + \int \Delta(\mu - \overline{\mu}) f(\overline{\mu})\ d\mu - \int \mu\Lambda(\mu) f'(\mu)\ d\mu + o(\delta) \\ &= 0\end{aligned}$$

Therefore, the intensive and extensive margins as defined in (N.95), are always zero: $\mathcal{I} = \mathcal{E} = 0$. The reason is that, on impact, a menu cost model with continuous idiosyncratic shocks does not generate first order effects.

**Application to our model** Now we apply the Caballero and Engel (2007) decomposition to our model, where the state is given by the markup estimate and its uncertainty $(\hat{\mu}, \Omega)$ and the inaction region is symmetric and two-dimensional $\mathcal{R} = [-\overline{\mu}(\Omega), \overline{\mu}(\Omega)]$. It follows that the aggregate price adjustment following a fully observable monetary shock (disclosed shock) $\delta > 0$ is given by:

$$\Delta p(\delta) = -\int_\Omega \int_{-\overline{\mu}(\Omega)}^{-\overline{\mu}(\Omega)+\delta} (\hat{\mu} - \delta) f(\hat{\mu}, \Omega)\ d\hat{\mu}\ d\Omega. \tag{N.96}$$

where we have substituted the adjustment hazard (equal to 1 for all firms that fall outside the inaction region). After the first order Taylor approximation around $\delta = 0$, we obtain:

$$\frac{\Delta p(\delta)}{\delta} = \left[\int_\Omega \overline{\mu}(\Omega) \underbrace{f(-\overline{\mu}(\Omega), \Omega)}_{=0} d\Omega + \int_\Omega \underbrace{[F(-\overline{\mu}(\Omega), \Omega) - F(-\overline{\mu}(\Omega), \Omega)]}_{=0} d\Omega\right] + o(\delta) = o(\delta), \tag{N.97}$$

where the first term $f(-\overline{\mu}(\Omega), \Omega) = 0$ is equal to zero due to regularity of the boundary of the continuation region (the probability that there is positive expected duration at the boundary is equal to zero). Again, we obtain that on impact there are no first order effects and therefore the intensive and extensive margins, as defined in (N.95), are both equal to zero. Similar outcomes are obtained when solving the model numerically in discrete time at the weekly frequency.

## N.2 Static Residual Decomposition

From the previous examples, we learn that focusing on first order effects on impact produce non-informative statistics in our model, since all the effects are second order. We now propose a decomposition that considers the extensive margin as a residual. First, we consider the a static response upon the impact of the aggregate shock and then generalize it to dynamic responses.

Let $f^\delta(\mu)$ be the density of the state after an aggregate shock of size $\delta$, $\mathbb{E}[\Delta p|\mu]$ be the expected size of a price change conditional on adjustment when the initial state is $\mu$, and $\Lambda(\mu)$ be the probability

[7]The Dirac delta function is defined formally as $\int_{-\infty}^{\infty} \Delta(x-a)f(x)dx = f(a)$, or informally as $\Delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}$.

[8]We use: $\Delta(x-a)f(x) = \Delta(x-a)f(a)$, $\Delta'(x)f(x) = -\Delta(x)f'(x)$, and $x^n\Delta(x) = 0\,\forall n > 0$.

of price change conditional on the state. Then, starting from the definition of inflation and doing a few manipulations:

$$\begin{aligned}\Delta p(\delta) &\equiv \int \mathbb{E}\left[\Delta p|\mu\right]\Lambda(\mu)f^{\delta}(\mu)\ d\mu \\ &= \int \delta\Lambda(\mu)f^{\delta}(\mu)d\mu + \int (\mathbb{E}\left[\Delta p|\mu\right]-\delta)\,\Lambda(\mu)f^{\delta}(\mu)d\mu \\ &= \delta\left[\int \Lambda(\mu)f^{\delta}(\mu)d\mu + \int\left(\frac{\mathbb{E}\left[\Delta p|\mu\right]-\delta}{\delta}\right)\Lambda(\mu)f^{\delta}(\mu)d\mu\right]\end{aligned}$$

we obtain the following decomposition, that computes the extensive margin as a residual:

$$\frac{\Delta p(\delta)}{\delta} = \mathcal{I}(\delta)+\mathcal{E}(\delta), \tag{N.98}$$

$$\mathcal{I}(\delta) \equiv \int \Lambda(\mu)f^{\delta}(\mu), \tag{N.99}$$

$$\mathcal{E}(\delta) \equiv \int\left(\frac{\mathbb{E}\left[\Delta p|\mu\right]-\delta}{\delta}\right)\Lambda(\mu)f^{\delta}(\mu). \tag{N.100}$$

As an example, consider the Calvo model with constant $\Lambda(\mu)=\Lambda$, which implies that

$$\mathcal{E}(\delta)=\Lambda\left[\int\left(\frac{\mathbb{E}\left[\Delta p|\mu\right]-\delta}{\delta}\right)f^{\delta}(\mu)d\mu\right] = \frac{\Lambda}{\delta}\left(\int \mathbb{E}\left[\Delta p|\mu\right]f^{\delta}(\mu)d\mu\right)-\Lambda = \frac{\Lambda}{\delta}\delta-\Lambda = 0$$

since the cross-sectional average of price changes is equal to the size of the money shock. Thus our decomposition respects the standard result that with Calvo pricing there is only an intensive margin of adjustment. Next we extend this new definition to compute dynamic responses.

## N.3 Dynamic Residual Decomposition (Disclosed Nominal Shock)

Now we decompose the impulse-response function into intensive and extensive margins. First, we will focus in the case with perfect observability of the nominal shock and then we analyze the case with partially disclosed monetary shock. From now on, we consider our model with information frictions, but the decomposition is analogous in other models.

Let $\mathbb{E}\left[\Delta p|\hat{\mu},\Omega\right]$ be the expected size of a price change conditional on adjustment when the initial state is $(\hat{\mu},\Omega)$, $\Lambda(\hat{\mu},\Omega)$ be the conditional probability of price change, and $f_t^{\delta}(\hat{\mu},\Omega)$ be the density of firms at time $t$. Let $\Delta p_t(\delta)$ be inflation at time $t$ following a monetary shock of size $\delta$, which can expressed as:

$$\Delta p_t(\delta) = \int_{\hat{\mu},\Omega} \mathbb{E}\left[\Delta p|\hat{\mu},\Omega\right]\ \Lambda(\hat{\mu},\Omega)\ f_t^{\delta}(\hat{\mu},\Omega)\ d\hat{\mu}d\Omega. \tag{N.101}$$

Since the nominal shock is disclosed (perfectly observed by all firms), once a firm adjusts its price it fully incorporates the nominal shock. As we discussed in the main text, only the first price change matters for the effects of nominal shocks; further price changes only reflect idiosyncratic shocks that wash out in the aggregate. For this reason, we can replace the density $f_t^{\delta}(\hat{\mu},\Omega)$ with the density conditional on inaction between $\{0,1,...,t\}$, that we denote with $f_t^{I,\delta}(\hat{\mu},\Omega)$, and only track the firms that have not adjusted yet. Our dynamic residual decomposition then becomes:

$$\frac{\Delta p_t(\delta)}{\delta} = \mathcal{I}_t(\delta)+\mathcal{E}_t(\delta), \tag{N.102}$$

$$\mathcal{I}_t(\delta) = \int \Lambda(\hat{\mu},\Omega)f_t^{I,\delta}(\hat{\mu},\Omega), \tag{N.103}$$

$$\mathcal{E}_t(\delta) = \int\left[\frac{\mathbb{E}\left[\Delta p|\hat{\mu},\Omega\right]-\delta}{\delta}\right]\Lambda(\hat{\mu},\Omega)f_t^{I,\delta}(\hat{\mu},\Omega)\ d\hat{\mu}d\Omega. \tag{N.104}$$

Notice that $\mathcal{I}_t^\delta$ satisfies the same definition as Caballero and Engel (2007) for the static impact effects, since in the first period the densities satisfy $f_0^{I,\delta}(\hat{\mu},\Omega,\delta) = f_0^\delta(\hat{\mu},\Omega,\delta)$. The difference lies in that this decomposition considers higher orders beyond the first. In the Calvo model, since $\Lambda(\hat{\mu},\Omega) = \Lambda$ and $f_t^{I,\delta}(\hat{\mu},\Omega) \propto f_0^{I,\delta}(\hat{\mu},\Omega)$, we have that $\Lambda \int \left[\frac{\mathbb{E}[\Delta p|\hat{\mu},\Omega]}{\delta}\right] f_t^{I,\delta}(\hat{\mu},\Omega,\delta) = 1$ and $\mathcal{E}_t^\delta(\Omega) = 0$ for all $t$.

## N.4 Dynamic Residual Decomposition (Partially Disclosed Nominal Shock)

Following the notation as in the main text, let $\alpha$ denote the observability of the nominal shock and $\varphi_0 = (1-\alpha)\delta$ be the biased component of forecast errors on impact. The case with a partially disclosed nominal shock ($\alpha < 1$) is challenging. The problem lies in that firms do not fully incorporate the nominal shock after the first price change (recall that in this case forecast errors arise), and therefore, inflation cannot longer be defined using $f_t^{I,\delta}(\hat{\mu},\Omega)$ as before:

$$\Delta p_t(\delta) \neq \int_{\hat{\mu},\Omega,\varphi} \mathbb{E}\left[\Delta p|\hat{\mu},\Omega,\varphi\right] \Lambda(\hat{\mu},\Omega,\varphi) f_t^{I,\delta}(\hat{\mu},\Omega,\varphi)\; d\hat{\mu} d\varphi d\Omega. \tag{N.105}$$

Our proposed solution is to replace the density $f_t^\delta(\hat{\mu},\Omega,\delta)$ with a new density $f_t^{II,\delta}(\hat{\mu},\Omega,\varphi)$, which keeps track of firms that have positive forecast errors or firms that have a zero forecast error but have not adjusted. In other words, the density $f_t^{II,\delta}(\hat{\mu},\Omega,\varphi)$ has dropped firms that satisfy the following two requirements:

- have a zero biased component of the forecast error $\varphi = 0$, and
- have changed the price between $\{0, 1, ..., t\}$.

Intuitively, if a firm has adjusted its price after she has fully updated its information set, then it means that it has fully incorporated the money shock into its price; consequently, additional price changes have a zero mean and can be ignored for computing the effects of a monetary shock. Therefore, we have that inflation can be computed as

$$\Delta p_t(\delta,\alpha) = \int_{\hat{\mu},\Omega,\varphi} \mathbb{E}\left[\Delta p|\hat{\mu},\Omega\right] \Lambda(\hat{\mu},\Omega) f_t^{II,\delta}(\hat{\mu},\Omega,\varphi) d\hat{\mu} d\Omega d\varphi. \tag{N.106}$$

Now we can decompose $\Delta p_t(\delta,\alpha)$ into the intensive margin and the extensive margin as before, but including a normalization factor that takes into account the observability $\alpha$:

$$\frac{\Delta p_t(\delta,\alpha)}{\delta} = \mathcal{I}_t(\delta,\alpha) + \mathcal{E}_t(\delta,\alpha), \tag{N.107}$$

$$\mathcal{I}_t(\delta,\alpha) = \int \Lambda(\hat{\mu},\Omega) \left(\frac{\delta-\varphi}{\delta}\right) f_t^{II,\delta}(\hat{\mu},\Omega,\varphi) d\hat{\mu} d\Omega d\varphi, \tag{N.108}$$

$$\mathcal{E}_t(\delta,\alpha) = \int \left[\frac{\mathbb{E}\left[\Delta p|\hat{\mu},\Omega\right] - (\delta-\varphi)}{\delta}\right] \Lambda(\hat{\mu},\Omega) f_t^{II,\delta}(\hat{\mu},\Omega,\varphi) d\hat{\mu} d\Omega d\varphi. \tag{N.109}$$

## N.5 Computation

Figure XXVI shows the inflation decomposition and the adjustment frequency for a disclosed monetary shock of size $\delta = 1\%$ together with an aggregate uncertainty shock of size $\kappa\mathbb{E}[\Omega]$, with $\kappa \in \{0, 1, 4\}$. We consider fully disclosed ($\alpha = 1$) and fully undisclosed shocks ($\alpha = 0$). We show statistics at the monthly frequency so that they are comparable to the data.

We can extract several conclusions from this exercise.

i) On impact, the intensive margin and the adjustment frequency always increase.

ii) Since the inflation change is larger than the intensive margin on impact (and in the first few months), the extensive margin also increases.

iii) As the aggregate uncertainty shock gets larger ($\kappa$ increases), the relative importance of the intensive margin increases on impact.

| On Impact: | No $\Omega$ shock $\kappa = 0$ | Small $\Omega$ shock $\kappa = 1$ | Large $\Omega$ shock $\kappa = 4$ |
|---|---|---|---|
| Intensive/Total | 30% | 37% | 45% |
| Extensive/Total | 70% | 63% | 55% |

iv) Without an aggregate uncertainty shock, the frequency becomes constant after the first months. The extensive margin disappears and the intensive margin is akin to that of the Calvo model. However, this property breaks with aggregate uncertainty shocks,

v) In the first months, the extensive margin is always positive, then it falls below zero, until it stabilizes. The reason is that after the monetary shock, a larger fraction of firms hit the lower boundary and increase their prices, while in the following months a larger fraction of firms hits the upper bound and lowers its prices.

Figure XXVI: Inflation Decomposition into Intensive and Extensive Margins Disclosed Shock

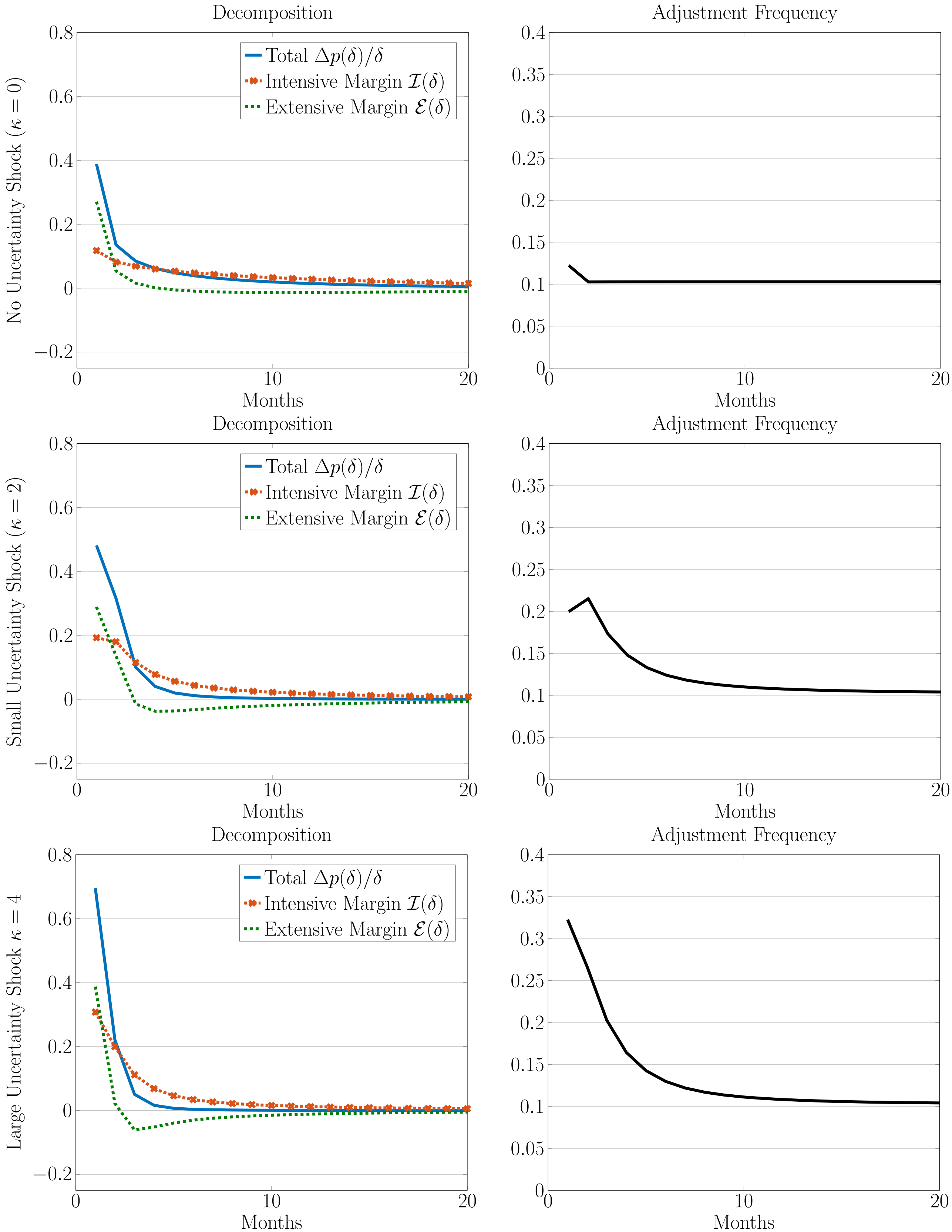
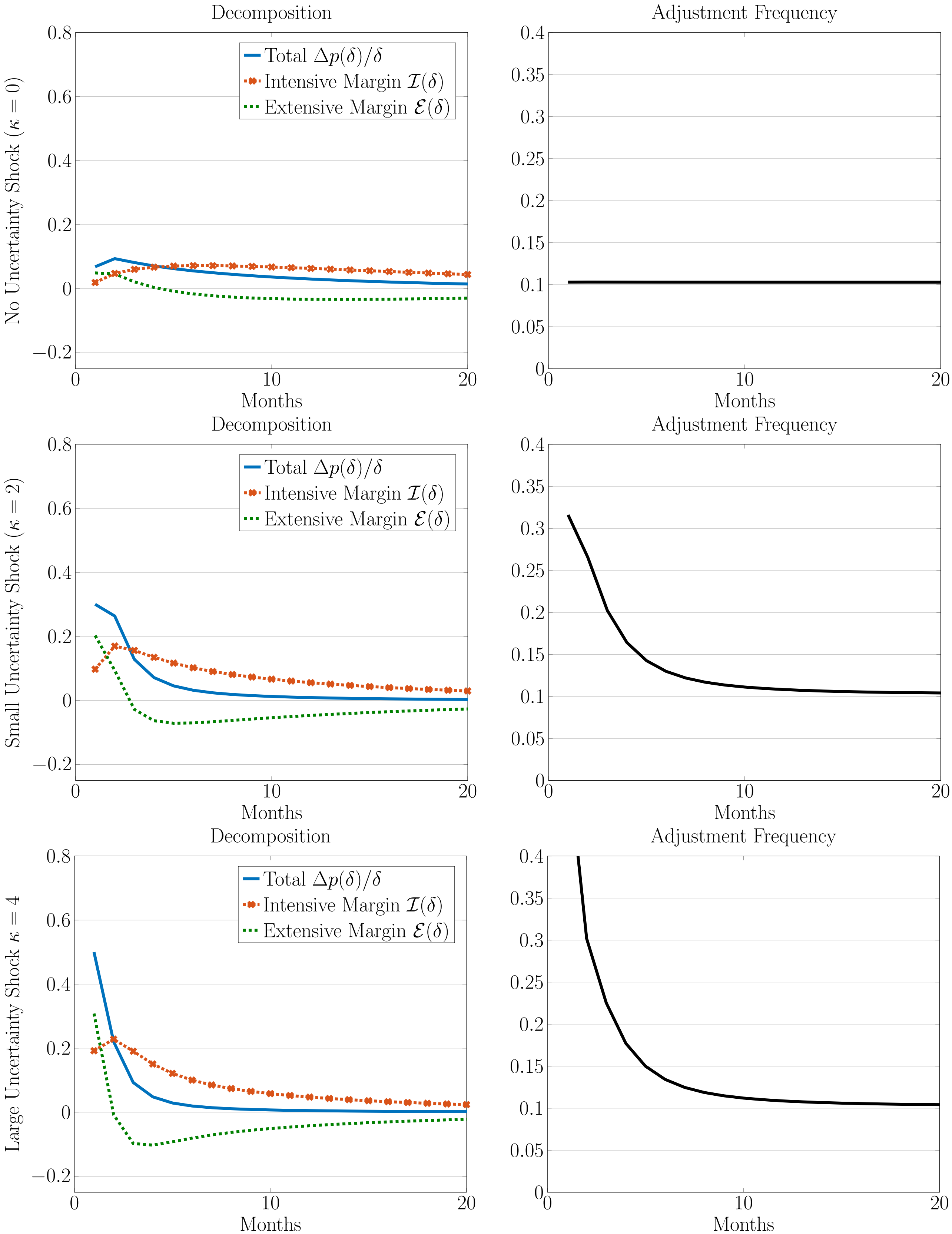


Figure XXVII: Inflation Decomposition into Intensive and Extensive Margins Undisclosed Shock



# O Uncertainty and pass-through

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

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

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

$$\Delta p_t^i = \beta_M \ \Delta_c M_t^i + \beta_\Omega \ \Omega_t^i + \beta_{M,\Omega} \ (\Delta_c M_t^i \times \Omega_t^i) + \epsilon_t^i \tag{O.110}$$

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

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

Uncertainty on its own is not statistically significant, but its interaction with the cumulative money shock yields very interesting results. When monetary shocks are observable in an environment of uncertain idiosyncratic productivity (Column 4b), the coefficient $\beta_{M,\Omega}$ is negative: when firm uncertainty is high, idiosyncratic productivity shocks become relatively more important than monetary shocks for pricing decisions; this reduces selection effects and average pass-through. In contrast, when the monetary shocks are unobservable (Column 5b), the coefficient $\beta_{M,\Omega}$ is positive. In this case, highly uncertain firms assign a larger Bayesian weight to observations that contain the monetary shock, and incorporate a larger fraction of the shock into their prices. Given the positive relationship between uncertainty and standard deviation of price changes in Proposition 8, our results imply a positive relationship between the standard deviation of price changes and pass-through, as documented in Berger and Vavra (2017) in the context of import price-setting.

The low pass-through of nominal shocks into individual prices is often attributed to strategic complementarities across firms. Our results show that information frictions about the nominal shock is an alternative way to decrease pass-through. Complementarities achieve it by decreasing the elasticity of

Table XII: Firm Uncertainty and Nominal Pass-Through

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

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

the size of price changes with respect to nominal marginal costs, whereas information frictions achieve it because firms take time to realize that costs have changed. However, the two mechanisms make opposite predictions regarding the relationship between idiosyncratic uncertainty and pass-through. For instance, Berger and Vavra (2017) shows that in model with strategic complementarities that arise from Kimball demand, larger volatility of idiosyncratic shocks reduces pass-through; while in our model with undisclosed nominal shocks, larger uncertainty about idiosyncratic shocks increases pass-through. This appears as an interesting implication for empirical research.

## P Additional Steps in Proof of Kalman-Bucy Filter with Jumps

In this section we show formally the convergence of the discrete time sampling observation. To show the convergence result we used Theorem 1.1 in Chapter 10 of Kushner and Dupuis (2013), where it shows weak convergence for Īto-Levy processes when the jumps are given by a compound Poisson process. To show this first we write the estimates as a function of the orthogonal innovation for the discrete sampling. We write the estimate in this way to apply the theorem in Kushner and Dupuis (2013). Then we show that the orthogonal innovation for the estimate is a Weiner process and for the variance is a Poisson process.

The following theorems are used in several of our proofs. These are standard so they are provided without proof.

**Theorem 2** *Fix a probability space $(\Lambda, \mathcal{F}, P)$. Let $X_j : \Lambda \to \mathcal{R}$ are Normal random variables. Then $X = (X_1, X_2, \ldots, X_n)$ is Normal if and only if $Y = \sum_i \lambda_i X_i$ is Normal.*

**Theorem 3** *Suppose that $X_k : \Lambda \to \mathcal{R}$ is normal for all $k$ and that $X_k \to X$ is $L^2$ . Then $X$ is normal*

**Theorem 4** *Let $B_t$ be a normal random variable for all $t$. Then $\int_0^t B_s ds$ is normal.*

**Theorem 5** *Let $Y, X|Z = z$ are normal random variables, then $Y|(X = x, Z = z)$ are normal random variables.*

**Theorem 6** *Let $y_t$ be a continuous process and $S$ a dense set in $[0, T]$. Then $\sigma\{y_s : s \in S\} = \sigma\{y_s : s \in [0, T]\}$*

**Theorem 7** *Let $Y : \Lambda \to \mathcal{R}$ be a random variable. Let $\mathcal{F}_n$ be an increasing sequence of $\sigma$-algebra. Then $X_n = Y|\mathcal{F}_n$ is a martingale with respect to $\mathcal{F}_n$*

**Theorem 8** *Let $M_k$ be a discrete time martingale with respect $\mathcal{F}_k$. Assume that $\sup_k \mathbb{E}[|M_k|^p] < \infty$ $for$ some $p$ Then there exist a $M \in L^1$ s.t. $M_t \to M$ a.e. $(P)$ and $\int |M_k - M| dP \to 0$ as $k \to \infty$.*

**Theorem 9** *Let $X \in L^1$, let $\mathcal{F}_k$ and increasing sequence of $\sigma$-algebra, $\mathcal{F}_k \in \mathcal{F}$ and define $\mathcal{F}_\infty = \sigma\{\cup_{i=1}^{\infty} \mathcal{F}_i\}$, then $\mathbb{E}[X|\mathcal{F}_k] \to \mathbb{E}[X|\mathcal{F}_\infty]$ a.e. in $P$ and in $L^1$.*

The following Lemma shows that the markup gap estimate can be written as a function of previous period estimate plus Normal orthogonal innovations.

**Lemma 3** *Let $t = 0, 1/2^n, 2/2^n, \ldots$ and $\Delta = \frac{1}{2^n}$. Define then $n-$estimate and $n-$variance as follows:*

$$\begin{aligned} \hat{\mu}_t^n &= \mathbb{E}[\mu_t | I_t^{n,*}] \\ \Sigma_t^n &= \mathbb{E}[(\mu_t - \hat{\mu}_t^n)^2 | I_t^{n,*}] \end{aligned}$$

*Then it follows that*

$$\begin{aligned} \hat{\mu}_{t+\Delta} - \hat{\mu}_t &= F \hat{\mu}_t \Delta + \varphi^I(\Delta) D X_{t+\Delta} \\ \Sigma_{t+\Delta} - \Sigma_t &= \left(\varphi^{II}(\Delta) - \sigma_f^2\right) \Delta + I(Q_{t+\Delta} - Q_t) \sigma_u^2 + o(\Delta^2) \end{aligned}$$

*where*

$$\begin{aligned} &\varphi^I(\Delta) \to_{n\to\infty} \frac{\Sigma_t G}{\gamma} \\ &\varphi^{II}(\Delta) \to_{n\to\infty} 2F\Sigma_t - \frac{G^2 \Sigma_t^2}{\gamma^2} \\ &D X_t \sim_{i.i.d} \mathcal{N}(o(\Delta), \Delta + o(\Delta^2)) \end{aligned}$$

**Proof.** From the main text we have that

$$\begin{aligned}\hat{\mu}^n_{t+\Delta} - \hat{\mu}^n_t &= \left(F - G\varphi^I(\Delta)\right)\hat{\mu}_t\Delta + \varphi^I(\Delta)\left(s_t - s_{t-\Delta}\right) + o(\Delta^2)\\ \Sigma^n_{t+\Delta} - \Sigma^n_t &= \left(\varphi^{II}(\Delta) + \sigma_f^2\right)\Delta + I(Q_{t+\Delta} - Q_t)\sigma_u^2 + o(\Delta^2)\end{aligned} \tag{P.111}$$

where

$$\begin{aligned}\varphi^I(\Delta) &= \frac{\Sigma_t G}{\Sigma_t G^2\Delta + \gamma^2} \rightarrow_{\Delta\to 0} \frac{\Sigma_t G}{\gamma^2}\\ \varphi^{II}(\Delta) &= (1+\Delta F)^2(1-K_tG\Delta)^2\Sigma_t + (1+\Delta F)^2K_t^2\gamma^2 - \Sigma_t \rightarrow_{\Delta\to 0} 2F\Sigma_t - \frac{G^2\Sigma_t^2}{\gamma^2}\end{aligned} \tag{P.112}$$

Given that

$$s_t = s_{t-\Delta} + G\Delta\mu_t + \sqrt{\Delta\gamma^2}\eta_t + o(\Delta^2) \tag{P.113}$$

and $\mathbb{E}[\mu_t - \hat{\mu}^n_t|\mathcal{I}^{n,*}_t] = 0$, we have that $\varphi^I(\Delta)\gamma\frac{s_t - s_{t-\Delta} - G\hat{\mu}_t\Delta}{\gamma} - o(\Delta^2) = DX_t$ where $DX_t$ is an i.i.d. Normal with mean $o(\Delta)$ and variance $\Delta + o(\Delta)$. Therefore

$$\hat{\mu}_{t+\Delta} - \hat{\mu}_t = F\hat{\mu}_t\Delta + \varphi^I(\Delta)DX_{t+\Delta}$$

with $DX_t \sim \mathcal{N}(o(\Delta), \Delta + o(\Delta))$ ■

To show convergence we extend the estimation for all $t$ to a right continuous process in the following way

$$(\hat{\mu}^n_t, \hat{\Sigma}^n_t) = (\hat{\mu}^n_{t_t}, \hat{\Sigma}^n_{t_t}) \quad \textit{with } t_i = \frac{k}{2^n} \textit{ and } t \in [t_i, t_i + \frac{1}{2^n}) \tag{P.114}$$

**Proposition 6** *Let $\xi^n_t = (\hat{\mu}^n_t, \hat{\Sigma}^n_t)$ be the solution of (P.111) and let $\vartheta_t = (\hat{\mu}_t, \Sigma_t)$ be the solution of*

$$\begin{aligned}d\hat{\mu}_t &= F\hat{\mu}_t dt + \frac{G^2\Sigma_t}{\gamma}dX_t\\ d\Sigma_t &= \left(2F\Sigma_t + \sigma_f^2 - \frac{G^2\Sigma_t^2}{\gamma^2}\right)dt + \sigma_u^2 dQ_t\end{aligned} \tag{P.115}$$

*The $\xi^n_t$ converge weekly to $\vartheta_t$.*

**Proof.** We follow the proof of Theorem 1.1, Chapter 10 of Kushner and Dupuis (2013). Let $\tau_n = \inf_t\{(\hat{\mu}^n_t, \Sigma^n_t) \in KK : KK \textit{ is compact}\}$ and define $\tilde{\xi}^n_t = \xi^n_{\min\{t,\tau^n\}}$. Notice that assumption A1.1 to A1.4 of Chapter 10 for $\tilde{\xi}$. The proof consists of 3 steps:

1. Show that $\xi^n_t$ is tight;
2. Show that the innovations are given by a Weiner or Poisson measure;
3. Show that $\xi^n$ weakly converges to $\vartheta_t$. To show convergence, we need to generate a bound on the drift and variance of the process.

**Tightness of the process $\tilde{\xi}^n_t$.** For all stopping times $\nu^n$ measurable with respect to the filtration generated by $\xi_t$, $\mathcal{F}^n_t$, and such that $\nu^n \le \tau^n$ we have that

$$\begin{aligned}&\mathbb{E}\left[\left((\hat{\mu}_{\nu^n+\Delta} - \hat{\mu}_{\nu^n})^2, (\Sigma_{\nu^n+\Delta} - \Sigma_{\nu^n})^2\right)|\mathcal{F}^n_t\right]\\ =\ &\mathbb{E}\left[\xi^n_{\nu^n+\Delta} - \xi^n_\nu|\mathcal{F}^n_t\right]^2 + \mathbb{E}\left[\left(\xi^n_{\nu^n+\Delta} - \xi^n_\nu - \mathbb{E}[\Delta\xi^n_{\nu^n}|\mathcal{F}^n_t]\right)^2|\mathcal{F}^n_t\right]\\ =\ &\left[(F\hat{\mu}^n_{\nu^n})^2, \left(2F\Sigma^n_{\nu^n} + \sigma_f^2 + \sigma_u^2\lambda - \frac{G^2(\Sigma^n_{\nu^n})^2}{\gamma^2}\right)^2\right]\Delta^2 + \left(\frac{G^2(\Sigma^n_{\nu^n})^2}{\gamma^2}, \sigma_u^4\lambda\right)\Delta + o(\Delta)\\ \le\ &(P_1, P_2)\Delta^2 + (P_3, \sigma_u^2\lambda)\Delta + o(\Delta)\\ =\ &O(\Delta)\end{aligned}$$

where we have defined

$$P_1 \equiv \max_{(x,y)\in KK} F^2 x^2 \qquad P_2 \equiv \max_{(x,y)\in KK} \left(2Fy + \sigma_f^2 + \sigma_u^2 \lambda - \frac{G^2 y^2}{\gamma^2}\right)^2 \qquad P_3 \equiv \max_{(x,y)\in KK} \frac{G^2 y^2}{\gamma^2}$$

Given that KK is compact and all the functions are continuous, the maximum exists.

**Innovations to markups gap estimates follow a Weiner process** Define the stochastic process

$$w_t^\Delta \equiv \sum_{i=0}^{n-1} \frac{\gamma}{G\Sigma_{t_i}^\Delta} \left(\Delta\hat{\mu}_{t_i} - \mathbb{E}_{\xi_{t_i}^\Delta}[\Delta\hat{\mu}_{t_i}^\Delta]\right) \qquad for\ t \in [t_n^\Delta, t_{n+1}^\Delta)$$

Then for all $f \in C^2(KK)$

$$\begin{aligned}
&= \quad f(w_{t+j}^\Delta) - f(w_t^\Delta) - \int_t^{t+j} 0.5 f''(w_s^\Delta) ds \\
&= \sum_{i=N_\Delta(t)}^{N_\Delta(t+j)-1} f'(w^h(t_i)) \frac{\gamma}{G\Sigma_{t_i}^\Delta} \left(\Delta\hat{\mu}_{t_i} - \mathbb{E}_{\xi_{t_i}^\Delta}[\Delta\hat{\mu}_{t_i}^\Delta]\right) + \\
&\dots \quad +0.5 \sum_{i=N_\Delta(t)}^{N_\Delta(t+j)-1} f''(w^h(t_i)) \frac{\gamma^2}{G^2(\Sigma_{t_i}^n)^2} \left(\Delta\hat{\mu}_{t_i} - \mathbb{E}_{\xi_{t_i}^\Delta}[\Delta\hat{\mu}_{t_i}^\Delta]\right)^2 - \dots \\
&\dots \quad -0.5 \sum_{i=N_\Delta(t)}^{N_\Delta(t+j)-1} f''(w^h(t_i))\Delta + \epsilon^\Delta + o(\Delta^2)
\end{aligned}$$

where $N_\Delta(t)$ is the number of jumps up to time $t$ and $\mathbb{E}[|\epsilon_t^\Delta|] \to 0$ as $\Delta \to 0$. Using the local consistency properties of the process, which are given by:

$$\begin{aligned}
\mathbb{E}_{\xi_t^\Delta}[\Delta\hat{\mu}_t^n] &= F\hat{\mu}_t^\Delta \Delta + o(\Delta) \\
\mathbb{E}_{\xi_t^\Delta}\left[\left(\Delta\hat{\mu}_t^n - \mathbb{E}_{\xi_t^n}[\Delta\hat{\mu}_t^\Delta]\right)^2\right] &= \frac{G^2\Sigma_t^2}{\gamma^2}\Delta + o(\Delta)
\end{aligned}$$

we have that for all function $H(\xi_{t_1}, ..., \xi_{t_q}, w_{t_1}, ..., w_{t_q})$ with $t_q \leq t$ bounded and continuous, the following holds:

$$\mathbb{E}[H(\cdot)\left(f(w_{t+j}^\Delta) - f(w_t^\Delta) - \int_t^{t+j} 0.5 f''(w_s^\Delta) ds\right)] \leq \mathbb{E}[\epsilon^\Delta] + o(\Delta^2) \to 0$$

From tightness we have that $w^n(\cdot)$ converges (see Chapter 9 of Kushner and Dupuis (2013) for the complete explanation of convergence) and the limit satisfies

$$\mathbb{E}[\left(f(w_{t+j}) - f(w_t) - \int_t^{t+j} 0.5 f''(w_s) ds\right)|\mathcal{F}_t] = 0$$

where $\mathcal{F}_t$ is the filtration generated by the limit random variable. If the limit random variable has continuous paths, then it is a Weiner process. Moreover, since

$$\frac{\mathbb{E}[|w_{t+\Delta}^n - w_t^n|^n]}{\Delta} = \frac{\left(\sqrt{\Delta} + o(\Delta)\right)^n}{\Delta} \to_{\Delta\to 0} 0$$

using the Kolmogorov Continuity Theorem we have continuity of the limit $w$. Therefore $w$ is a Wiener process.

**Innovations to estimate's variance follow a Poisson process** Define the stochastic process

$$Q_t^n \equiv \sum_{i=0}^{n-1} \left( \Delta\hat{\Sigma}_{t_i}^n - \mathbb{E}_{\xi_{t_i}^\Delta}[\Delta\hat{\Sigma}_{t_i}^n] \right)$$

Following similar steps as before, and using the local consistency conditions for $\Sigma_t^n$ we get that

$$\begin{aligned} \mathbb{E}_{\xi_t^\Delta}[\Delta\Sigma_t^n] &= \left( 2F\Sigma_t + \sigma_f^2 + \sigma_u^2\lambda - \frac{G^2\Sigma_t^2}{\gamma^2} \right) \Delta \\ \mathbb{E}_{\xi_t^\Delta}\left[ \left( \Delta\Sigma_t^\Delta - \mathbb{E}_{\Sigma_t^\Delta}[\Delta\Sigma_t^\Delta] \right)^2 \right] &= \sigma_u^4\lambda\Delta + o(\Delta) \end{aligned}$$

Then for all $f \in C^2(KK)$ with support $KK$ and for all function $H(\xi_{t_1}, ..., \xi_{t_q}, w_{t_1}, ..., w_{t_q})$ with $t_q \leq t$ bounded and continuous,

$$\mathbb{E}[H(\cdot) \left( f(Q_{t+j}^n) - f(Q_t^n) \right)] \to \mathbb{E}[H(\cdot)\lambda \int_t^{t+j} \left( f(Q_s + \sigma_u^2) - f(Q_s) \right) ds] = \mathbb{E}[H(\cdot) \left( f(Q_{t+j}) - f(Q_t) \right)]$$

As before $Q_t$ the limit of $Q_t^n$ is a Poisson measure. The rest of the proof is identical to Kushner and Dupuis (2013). ■

**Proposition 7** *Let $Y_t = (\hat{\mu}_t^Y, \Sigma_t^Y)$ be the solution of (P.115) and let $M_t = (\hat{\mu}_t^M, \Sigma_t^M)$ be the solution of*

$$\begin{aligned} d\hat{\mu}_t &= \left( F - \tfrac{G^2\Sigma_t}{\gamma^2} \right) \hat{\mu}_t dt + \tfrac{G\Sigma_t}{\gamma^2} ds_t \\ d\Sigma_t &= \left( 2F\Sigma_t + \sigma_f^2 - \tfrac{G^2\Sigma_t^2}{\gamma^2} \right) dt + \sigma_u^2 dQ_t \end{aligned} \tag{P.116}$$

*Then $M_t$ coincides in law with $Y_t$.*

**Proof.** Note that the law of motion for $\Sigma_t$ is the same, so trivially they coincide in law. For $\hat{\mu}_t$, notice that we can write P.115

$$\begin{aligned} d\hat{\mu}_t^Y &= F\hat{\mu}_t^Y dt + \frac{G\Sigma_t}{\gamma} dX_t \\ &= F\hat{\mu}_t^Y dt + \frac{G\Sigma_t}{\gamma} \left( \frac{G}{\gamma} \left( \mu_t - \mu_t^Y \right) dt + dH_t \right) \end{aligned} \tag{P.117}$$

where $H_t$ is another Brownian motion. The second equality is true, since by the Law of Iterated Expectations

$$\mathbb{E}\left[ (\mu_t - \hat{\mu}_t^Y) | \sigma\{\mu_s^Y : s \leq t\} \right] = \mathbb{E}\left[ \mathbb{E}\left[ (\mu_t - \hat{\mu}_t^Y) | \mathcal{I}_t \right] | \sigma\{\mu_s^Y : s \leq t\} \right] = \mathbb{E}[0 | \sigma\{\mu_s^Y : s \leq t\}] = 0 \tag{P.118}$$

where in the last step we used that $\hat{\mu}_t^Y$ is an unbiased estimator. Given that $\mathbb{E}[(\mu_t - \hat{\mu}_t^Y) | \sigma\{\mu_s^Y : s \leq t\}] = 0$, using corollary 8.4.5 of Øksendal (2007) we have that $X_t$ is a brownian motion if and only if $H_t$ is a brownian. Rewriting equation P.117, we have that

$$d\hat{\mu}_t^Y = \left( F - \frac{G^2\Sigma_t}{\gamma^2} \right) \hat{\mu}_t^Y dt + \frac{G\Sigma_t}{\gamma^2} \left( G\mu dt + \gamma dH_t \right) \tag{P.119}$$

Notice that if $H_t = Z_t$, then we have the result, since $G\mu dt + \gamma dH_t = ds_t$ ■

## References


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

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

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

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

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

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

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

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

Golosov, M. and Lucas, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, **115** (2), 171–199.

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

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

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

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

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

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

Kushner, H. J. and Dupuis, P. G. (2013). *Numerical Methods for Stochastic Control Problems in Continuous Time.*, vol. 24. Springer Science & Business Media.

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

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

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

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

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

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

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

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