

Online Appendix For: Pricing Mistakes and the Non-Neutrality of Money

Tomás E. Caravello
MIT

October 28, 2026

[\[Link to latest version\]](#)[\[Link to main text\]](#)

This appendix contains supplemental material for the article "Pricing Mistakes and the Non-Neutrality of Money." It provides: (i) additional empirical results; (ii) additional theoretical and quantitative results; (iii) supplementary proofs; (iv) general equilibrium microfoundations and details of the approximations; and (v) numerical solution procedures.

References to equations, figures, tables, assumptions, propositions, lemmas, or sections without the prefix "OA." refer to the main article, including its attached appendix.

Contents

OA.A	Additional Empirical Results	3
OA.A.1	Extensive Margin	3
OA.A.2	Small and Large Chains	4
OA.A.3	Correlation Between Frequency and Size of Price Changes	5
OA.B	Additional Theoretical and Quantitative Results	5
OA.B.1	A Deeper Dive Into Existing Signal Extraction Models	5
OA.B.2	Two-Layer Decision Process and Full Weight on Signals	9
OA.B.3	Endogenous Reset Distribution	12
OA.B.4	Quality of the Approximation	14
OA.B.5	Extended Identification Result	16
OA.B.6	Additional Discussion on the Structural Model	22
OA.B.7	Examples for Proposition 9	25
OA.C	Additional Proofs	29
OA.C.1	Propositions OA.1 and OA.2	29
OA.C.2	Proof of Proposition 12	35
OA.C.3	Invariance to Small Inflation Trend	38
OA.C.4	Proof of Calvo Hazard	40
OA.C.5	Calvo Model with Mistakes and Mean Reversion	41
OA.D	Microfoundations and Approximations	52
OA.D.1	General Equilibrium Setup	52
OA.D.2	Continuous Time as a Limit of Discrete Time	56
OA.E	Numerical Solution Procedure	57
OA.E.1	One Variable	58
OA.E.2	Two Variables	60

Time	Vintage	$\Delta p_{i,n-1}$	$\Delta p_{i,n}$	$\mathbf{1}\{\tau_{i,n} = s \wedge \Delta p_{i,n} > 0\}$	$\mathbf{1}\{\tau_{i,n} = s \wedge \Delta p_{i,n} < 0\}$
t-1	-	-	-	0	0
t	-	-	0.1	1	0
t+1	t	0.1	-	0	0
t+2	t	0.1	-	0	0
t+3	t	0.1	0.02	1	0
t+4	$t + 3$	0.02	-	0	0
t+5	$t + 3$	0.02	-0.05	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table OA.1: Example Dataset. Here, a product first changes prices at t by 0.1, and then again at $t + 3$, by 0.02, and once more at $t + 5$, by -0.05 .

OA.A Additional Empirical Results

OA.A.1 Extensive Margin

This section contains results for the extensive margin. In particular, we consider the following two regression specifications:

$$\mathbf{1}\{\tau_{i,n} = s \wedge \Delta p_{i,n} > 0\} = \text{Fixed Effects} + \beta_u \Delta p_{i,n-1} + \text{Error}, \quad (\text{OA.A.1})$$

$$\mathbf{1}\{\tau_{i,n} = s \wedge \Delta p_{i,n} < 0\} = \text{Fixed Effects} + \beta_d \Delta p_{i,n-1} + \text{Error}, \quad (\text{OA.A.2})$$

where $\mathbf{1}\{\tau_{i,n} = s \wedge \Delta p_{i,n} > 0\}$ is a dummy variable that equals 1 if the price increases at time s and 0 otherwise. The second specification is the same, but for price cuts. Importantly, we run the regression for all s until there is a price change. For example, suppose that a product that has been observed since $t - 1$ first changes prices at t by 0.1, and then again at $t + 3$, by 0.02, and once more at $t + 5$ by -0.05 . Table OA.1 shows what the dataset would look like. Essentially, the specifications (OA.A.1)-(OA.A.2) estimate how much the probability of the price increasing (decreasing) in the next period changes based on the size of the last price change.

Table OA.2 presents the results for the UK, and Figure OA.A.1 shows the results for Argentina. As we can see, the patterns are intuitively consistent with the reversals in sizes we documented above: price increases predict a lower probability of increasing the price again, and a higher probability of cutting the price. For the UK, there are also significant differences across categories by frequency: coefficients are larger in magnitude when the

Sales Filter	β_u			β_d		
	Low	High	Difference	Low	High	Difference
V-Shaped	-0.099*** (0.015)	-0.219*** (0.012)	-0.120*** (0.019)	0.139*** (0.021)	0.305*** (0.013)	0.166*** (0.025)
NS(A)	-0.068*** (0.011)	-0.198*** (0.015)	-0.130*** (0.019)	0.152*** (0.023)	0.345*** (0.015)	0.194*** (0.028)
KM	-0.111*** (0.011)	-0.148*** (0.005)	-0.037*** (0.012)	0.100*** (0.010)	0.150*** (0.006)	0.050*** (0.011)

Table OA.2: Estimates of predictability regressions (OA.A.1)-(OA.A.2) in different subsamples. All regressions include time-by-cell fixed effects. “Low” corresponds to categories with low frequency of price changes (below the average frequency of one price change every 9.23 months), whereas “high” means categories with high frequency. Each row represents a different treatment of sales: row 1 uses a V-shaped filter, row 2 uses Filter A in Nakamura and Steinsson (2008), row 3 uses the Kehoe and Midrigan (2015) filter. Standard errors are clustered by item.

frequency is higher.

For the Argentine case, the effect on the probability of increasing is more sensitive during the Milei period, but the probability of cutting prices is less sensitive. Given that inflation interacts with the probability of price changes in a non-linear way, it is less clear what should happen with predictability coefficients: when inflation is as high as 25% per month, there are essentially no price cuts, so affecting their probability is much harder. The fact that I find an effect on the probability of price cuts for a period of such high inflation is remarkable.

OA.A.2 Small and Large Chains

The UK database includes a classification of shop types in which the product is sold. Each shop belongs to one of two categories: “Multiple” represents chains with ten or more outlets, whereas “Independent” represents chains with nine or fewer outlets. Table OA.3 presents the results when we run (3.1) splitting the sample by shop type. As we can see, no clear pattern emerges: for the V-shaped filter the difference is not significant, whereas the other filters show significant results, but in opposite directions. Overall, these results show that the size of the chain does not correlate with the strength of reversals in a clear way.

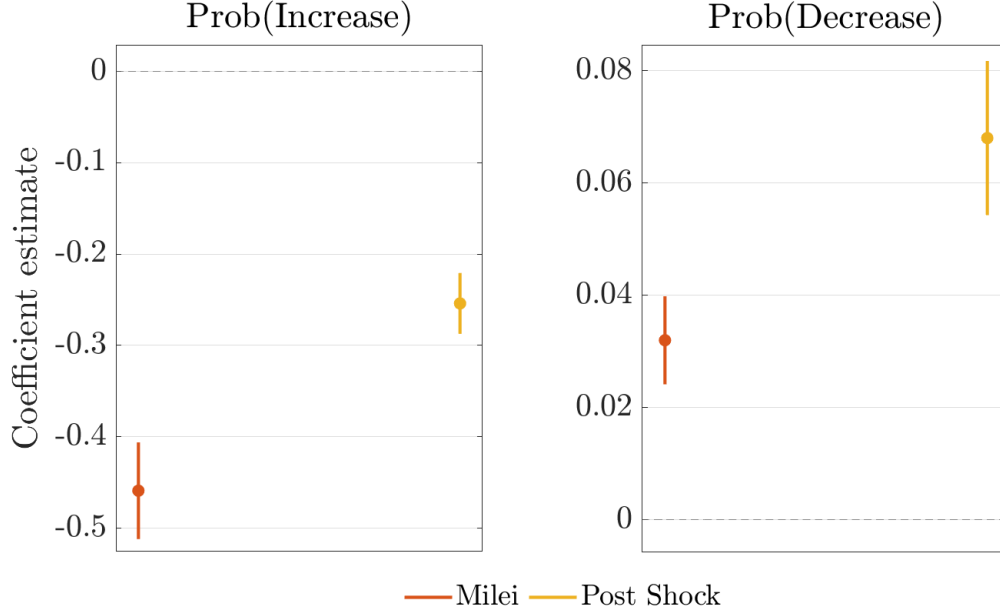


Figure OA.A.1: Estimates of extensive margin regression coefficients (OA.A.1) and (OA.A.2) for Argentina.

OA.A.3 Correlation Between Frequency and Size of Price Changes

Figure OA.A.2 presents the scatter plot of the average size of price changes and frequency, computed for each item in the UK sample. The red line corresponds to a fitted quadratic. As we can see, the relation is non-monotone, and the R^2 is very low. This indicates that the relation between these two variables is quite weak.

OA.B Additional Theoretical and Quantitative Results

OA.B.1 A Deeper Dive Into Existing Signal Extraction Models

Consider the setup in Baley and Blanco (2019). The true price gap x follows $dx = \sigma_f dW + \sigma_u u_t dQ_t$, where $u_t Q_t$ is a compound Poisson process, where u_t is normal and Q_t is a counting process. Firms do not observe the gap directly, but observe continuous signals s_t , which satisfy $ds_t = x dt + \gamma dZ$, where dZ is noise. Assuming profits are quadratic and overall Gaussianity, the firm has two state variables: its perceived gap $\hat{x}_t = \mathbb{E}[x_t | I_t]$, and its uncertainty about it, $\Omega_t = \mathbb{E}[(x_t - \hat{x}_t)^2 | I_t] / \gamma$, where I_t is the information set of the firm. Baley

Sales Filter	β_1		
	Multiple	Independent	Difference
No filter	-0.745*** (0.013)	-0.504*** (0.011)	0.240*** (0.017)
Flag	-0.505*** (0.012)	-0.450*** (0.009)	0.055*** (0.015)
V-Shaped	-0.347*** (0.010)	-0.364*** (0.014)	-0.017 (0.017)
NS(A)	-0.327*** (0.010)	-0.373*** (0.014)	-0.047*** (0.017)
KM	-0.362*** (0.020)	-0.297*** (0.029)	0.065* (0.035)

Table OA.3: Estimates of predictability regressions in (3.1) in different subsamples. Per ONS classification, “Multiple” represents chains with ten or more outlets, whereas “Independent” represents chains with nine or fewer outlets. Standard errors are clustered by item.

and Blanco (2019) show that $d\hat{x} = \sqrt{\Omega_t}d\hat{Z}_t$, where $\hat{Z}_t$ is a Wiener process under the firm’s information set. Given symmetry of the setup, it is optimal for firms to close their perceived price gaps in expectation, so price changes satisfy $p_{\tau_{i,n}} - p_{\tau_{i,n-1}} = -\hat{x}_{\tau_{i,n}}$. Consider now the covariance between two consecutive price changes, where for simplicity we condition on $\tau_{i,n-1} = t$, i.e., the first price change occurs at time t for all firms being compared. Given that price gaps have mean zero, the covariance between consecutive price changes is thus $\text{Cov}(\Delta p_{i,n-1}, \Delta p_{i,n}) = \mathbb{E}[\hat{x}_t \hat{x}_{\tau_{i,n}}]$. But $\mathbb{E}[\hat{x}_t \hat{x}_{\tau_{i,n}}] = \mathbb{E}[\hat{x}_t \mathbb{E}[\hat{x}_{\tau_{i,n}} | I_t]] = 0$: since $\hat{x}_t$ follows a Brownian process under the information set of the firm, it is a martingale, and since it starts at zero after reset, its conditional mean is zero. Thus, this model implies zero predictability.

While quite general, the proof above does not make the mechanism explicit when considering standard signal-extraction problems. In order to shed light on this, consider a stylized setup similar to Afrouzi et al. (2026). Assume a time-dependent model where we reset deterministically every T periods as in Taylor (1979) and everything is Gaussian. Assume the pricer has a prior $p_t^* \sim \mathcal{N}(\hat{p}_t, U_t)$, and each time we change prices, we get a signal $s = p^* + \sigma_u u$.¹ Assume that p^* follows a Brownian motion, and the noise u is independent. Assume that when an opportunity to adjust prices arrives, the pricer sets $p_t = \hat{p}_t$, i.e., it fully closes its price gap. As usual, this is optimal if the profit function is symmetric and there is no trend inflation.

¹This is for analytical convenience. A similar point would apply if we continuously observed signals as in Baley and Blanco (2019).

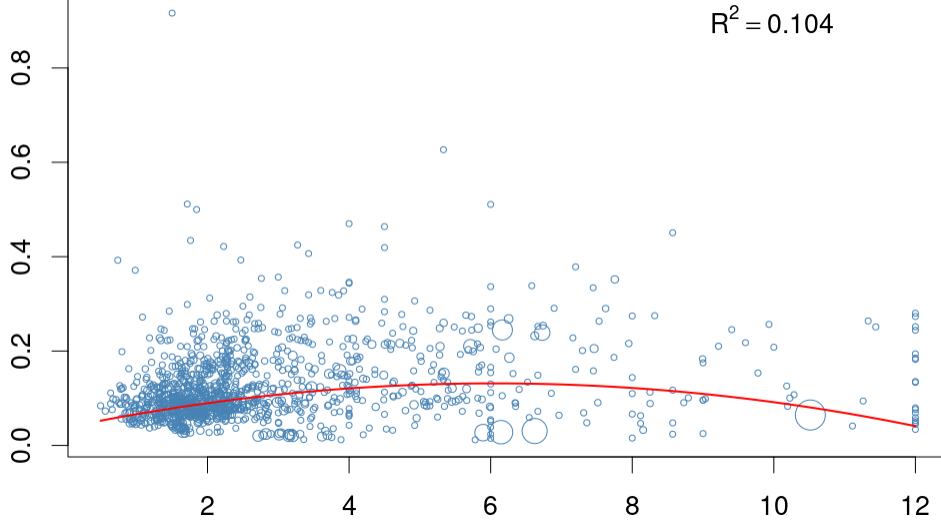


Figure OA.A.2: Scatter plot of the frequency of price changes (number per year, horizontal axis) and their average absolute size $|\Delta p|$ (vertical axis). Each dot represents an item; the circle's size indicates the number of price changes in that item. The red line shows the fit of the regression $|\Delta p|_i = \gamma_0 + \gamma_1 N_i + \gamma_2 (N_i)^2 + u_i$. The figure also reports the corresponding R^2 .

Assume that we are in a stationary environment, and that the best guess for the mean of p_t^* is updated according to

$$\hat{p}_{t+T} = K\hat{p}_t + (1 - K)s_{t+T}.$$

I study what happens with predictability for arbitrary values of K . Note that optimal Bayesian inference for an agent that tracks $\hat{p}$ features:

$$K = \frac{1/U_{t+T}}{1/U_{t+T} + (1/\sigma_u^2)}.$$

Let's derive the stationary values of uncertainty. First, note that after reset, uncertainty starts at $\underline{U}$ and right before reset rises to $\bar{U} = \underline{U} + \sigma^2 T$, where σ^2 is the variance of the fundamental. In order to find $\underline{U}$, note that due to the Bayesian updating, $1/\underline{U} = \frac{1}{\bar{U}} + \frac{1}{\sigma_u^2}$ so $\underline{U} = \frac{\bar{U}\sigma_u^2}{\bar{U} + \sigma_u^2}$. Putting both equations together we obtain $P(\bar{U}) = 0$, where $P(x) = x^2 - \sigma^2 T x - \sigma_u^2 \sigma^2 T$. This has two roots, $\bar{U} = \frac{\sigma^2 T \pm \sqrt{(\sigma^2 T)^2 + 4\sigma_u^2 \sigma^2 T}}{2}$. I take the positive one, $\bar{U} = \frac{\sigma^2 T + \sqrt{(\sigma^2 T)^2 + 4\sigma_u^2 \sigma^2 T}}{2}$, and thus $\underline{U} = \frac{-\sigma^2 T + \sqrt{(\sigma^2 T)^2 + 4\sigma_u^2 \sigma^2 T}}{2}$. Iterating

the updating equation backwards:

$$\hat{p}_{t+T} = (1-K) \sum_{\ell=0}^{\infty} K^{\ell} s_{t+T(1-\ell)},$$

and so:

$$\hat{p}_{t+T} - \hat{p}_t = (1-K) \sum_{\ell=0}^{\infty} K^{\ell} (s_{t+T(1-\ell)} - s_{t+T(-\ell)}),$$

but $s_{t+T(1-\ell)} - s_{t+T(-\ell)} = p_{t+T(1-\ell)}^* - p_{t+T(-\ell)}^* + \sigma_u(u_{t+T(1-\ell)} - u_{t+T(-\ell)})$ and since $p_{t+T(1-\ell)}^* - p_{t+T(-\ell)}^* = \sigma(W_{t+T(1-\ell)} - W_{t+T(-\ell)})$, this has a normal distribution with variance $\sigma^2 T$ (and the increments are independent of each other). Let $Z_{t+T(1-\ell)} = W_{t+T(1-\ell)} - W_{t+T(-\ell)}$ be a normal random variable with variance T . Thus we can write:

$$\begin{aligned} \hat{p}_{t+T} - \hat{p}_t &= (1-K) \sum_{\ell=0}^{\infty} K^{\ell} (\sigma Z_{t+T(1-\ell)} + \sigma_u(u_{t+T(1-\ell)} - u_{t+T(-\ell)})) \\ &= (1-K) \sigma \sum_{\ell=0}^{\infty} K^{\ell} Z_{t+T(1-\ell)} + (1-K) \sigma_u \left(u_{t+T} - (1-K) \sum_{\ell=0}^{\infty} K^{\ell} u_{t-T\ell} \right), \end{aligned}$$

which can also be written as $\hat{p}_{t+T} - \hat{p}_t = K(\hat{p}_t - \hat{p}_{t-T}) + (1-K)(\sigma Z_{t+T} + \sigma_u(u_{t+T} - u_t))$ and thus

$$\text{Cov}(\hat{p}_{t+T} - \hat{p}_t, \hat{p}_t - \hat{p}_{t-T}) = K \text{Var}(\hat{p}_t - \hat{p}_{t-T}) + (1-K) \text{Cov}(\sigma Z_{t+T} + \sigma_u(u_{t+T} - u_t), \hat{p}_t - \hat{p}_{t-T}),$$

note that:

$$\begin{aligned} \text{Var}(\hat{p}_t - \hat{p}_{t-T}) &= (1-K)^2 \left[\frac{\sigma^2 T}{1-K^2} + \sigma_u^2 \left(1 + \frac{(1-K)^2}{1-K^2} \right) \right] \\ &= \frac{(1-K)}{1+K} [\sigma^2 T + 2\sigma_u^2(1-K)], \end{aligned}$$

and

$$\begin{aligned} \text{Cov}(\sigma Z_{t+T} + \sigma_u(u_{t+T} - u_t), \hat{p}_t - \hat{p}_{t-T}) &= \text{Cov}(\sigma_u(-u_t), \hat{p}_t - \hat{p}_{t-T}) \\ &= -(1-K) \sigma_u^2, \end{aligned}$$

so

$$\begin{aligned}\text{Cov}(\hat{p}_{t+T} - \hat{p}_t, \hat{p}_t - \hat{p}_{t-T}) &= K \frac{(1-K)}{1+K} [\sigma_u^2 T + 2\sigma_u^2(1-K)] - (1-K)^2 \sigma_u^2 \\ &= \frac{1-K}{1+K} (K\sigma_u^2 T - (1-K)^2 \sigma_u^2).\end{aligned}\tag{OA.B.1}$$

This expression shows the two effects of K : the part with the fundamental shocks generates positive covariance due to under-reaction, whereas the part with noise generates negative correlation due to correction of mistakes. Evaluating this at the Bayesian value $K = \frac{\sigma_u^2}{\bar{U} + \sigma_u^2}$, we get

$$\begin{aligned}\frac{\sigma_u^2}{\bar{U} + \sigma_u^2} \sigma_u^2 T &= (1 - \frac{\sigma_u^2}{\bar{U} + \sigma_u^2})^2 \sigma_u^2 \\ (\bar{U} + \sigma_u^2) \sigma_u^2 T &= \bar{U}^2.\end{aligned}$$

This equality follows from the equation determining $\bar{U}$. Thus, the covariance in Equation (OA.B.1) is zero, and Bayesian updating implies zero predictability. This implies that in a Bayesian learning model, if information is used optimally, there will be no predictability. Note that this holds for any variance of the signal, so it does not matter whether σ_u is endogenous or exogenous.

Using equation (OA.B.1), we can also see what would happen if K was set to a different value. If $K > \sigma_u^2/(\bar{U} + \sigma_u^2)$, the pricer would underweight the signal. This underreaction would lead to updating that is too slow, and thus the covariance of consecutive price changes would be positive. In contrast, if $K < \sigma_u^2/(\bar{U} + \sigma_u^2)$, the pricer would overweight the signal. This would give too much weight to noise, which would be subsequently corrected and thus would generate reversals. I show in Section OA.B.2 below that a possible microfoundation for the control cost model in Section 5.2 is obtained if we assume $K = 0$, so new prices are obtained by putting full weight on new signals.

OA.B.2 Two-Layer Decision Process and Full Weight on Signals

Consider the general model in Section 5.2. Assume that the decision on when to change prices is taken by a *general manager* that can perfectly observe aggregate profits but does not know the exact price gaps of each product. This general manager delegates the decision of choosing each individual price to *pricing managers*. Apart from the order to update prices, no information is shared from the general manager to the pricing manager. As shown before,

the problem of the *general manager* can be characterized by

$$0 = \min \left\{ \rho V - (-By + n\sigma^2 V_y + 2\sigma^2 y V_{yy} + \kappa \int_0^\Psi \max\{V^*(0) - (\psi + V(y)), 0\} dG(\psi)), V(y) - V^*(\Psi) \right\},$$

When the general manager chooses to reset, she delegates the decision of choosing new prices to *pricing managers*, whose optimization problem I describe shortly. She correctly anticipates that, after reset, she will get an expected value $V^*(0)$.

Turning to the *pricing manager*, they choose the new price to maximize the profits of a given product line. Define:

$$V^p(x) = -B\mathbb{E}\left[\int_0^\infty e^{-\rho t} x_t^2 dt | x(0) = x\right],$$

where the expectation is taken both over time, and over the pricing decisions of other managers. Letting $x = p - p^*$, we can also write this as $V^p(p - p^*)$. Assume, at the time of reset, the manager has a symmetric prior about p^* . However, before resetting, the manager observes a signal $s = p^* + x^*$, where x^* has some distribution with density f . Assume that x^* is mean zero and symmetric. Using only the information conveyed by the signal, the manager has to choose the optimal price p .² First, note that:

$$p^{\max} = \arg \max_p \mathbb{E}[V^p(p - p^*) | s],$$

and the FOC is

$$\mathbb{E}[V_x^p(p - p^*) | s] = 0.$$

Now, note that if $p = \mathbb{E}[p^* | s]$, given that both the prior and the signal are symmetric, the posterior density for p^* is symmetric. So $p - p^* | s$ is a symmetric distribution with mean zero. Given that $V^p(x)$ is symmetric, V_x^p is antisymmetric, and thus $\mathbb{E}[V_x^p(\mathbb{E}[p^* | s] - p^*) | s] = 0$, so $p^{\max} = \mathbb{E}[p^* | s]$. Given that, we need to characterize $\mathbb{E}[p^* | s]$. Using Bayes's rule $f(p^* | s) \propto f(s | p^*) f(p^*)$, but $f(s | p^*) = f(s - p^*)$. Assuming that the prior $f(p^*)$ is uniform with increasingly large bounds (i.e., the pricing manager has a totally diffuse prior on p^*), we get $f(p^* | s) = f(s - p^*)$, and thus the conditional distribution of $p^* | s$ is the same distribution

²I assume that the pricing manager does not use the fact that it was requested to reset its price to infer the new value. While this assumption may sound extreme in the case with small n , when n is large the idea is that the order to reset prices does not provide meaningful information about the reset gap of each individual product.

of x^* , but shifted by p^* . And thus

$$\begin{aligned}\mathbb{E}[p^*|s] &= \int p^* f(s - p^*) dp^* \\ &= s + \int z f(z) dz = s,\end{aligned}$$

where I used symmetry in f and used the change of variables $z = p^* - s$. Intuitively, since the firm has a very flat prior on p^* , its posterior mean equals the signal, which is noisy. Thus the pricing manager chooses $p^{\max} = s = p^* + x^*$, or written in terms of price gaps, $p^{\max} - p^* = x^*$.

Now assume that the general manager can choose to expend effort to reduce the uncertainty of the signal faced by the pricing managers. Intuitively, this would mean that the general manager calls for more meetings or monitors the pricing managers more closely. In particular, she can make the pricing managers have a distribution f by paying a cost $\mathcal{C}(f)$ per product. The problem of choosing the distribution optimally is thus:

$$V^*(0) = \max_{\{f_i^y\}_{i=1}^n} \int \dots \int V\left(\sum_{i=1}^n y_i\right) \prod_{i=1}^n f_i^y(y_i) dy_i - \sum_{i=1}^n \mathcal{C}(f_i^y) \text{ s.t. } \int_0^\infty f_i^y(y_i) dy_i = 1 \quad \forall i,$$

which is exactly the control cost problem we have been working with.

DISCUSSION. The above justification formalizes the intuitive explanation given in Costain and Nakov (2015) by explicitly considering the pricing process as driven by different agents within the firm with different incentives and information. In this case, control costs are interpreted as monitoring costs that a general manager can pay to make their prices be more accurate.

Furthermore, in conjunction with the discussion in Appendix OA.B.1, it also directly connects with the imperfect information literature. Relative to noisy information models, the key departure is that no single agent is continuously tracking the desired price. Instead, agents choosing prices give full weight to signals, which introduces noise into prices, which is subsequently corrected. A corollary of this is that in the present paper, firms are fully aware of changes in aggregate variables. Thus, any implications for monetary non-neutrality are *not* driven by aggregate-idiosyncratic confusion as in the classical Lucas (1972) islands model, or by firms slowly learning about changes in aggregate states, as in Mankiw and Reis (2002); Sims (2003); Maćkowiak and Wiederholt (2009). Relative to the infrequent information literature (Reis, 2006; Woodford, 2009; Alvarez et al., 2011, 2016b), the true

value of the desired price is not perfectly observed at the moment of reset, and thus some noise leaks into the price, which generates future reversals.

OA.B.3 Endogenous Reset Distribution

Consider the following specification of the price adjustment cost functional: $\mathcal{C}(F) = \omega \hat{c}(\theta) + \theta(1 - \omega) \mathcal{I}(F || \bar{F})$. In this case, $\mathcal{C}(F)$ has two components: first, a fixed cost of size $\omega \hat{c}(\theta)$; second, a control cost $\theta(1 - \omega) \mathcal{I}(F || \bar{F})$, where as before $\mathcal{I}(F || \bar{F}) = \int \log \left(\frac{dF(x)}{d\bar{F}(x)} \right) dF(x)$ is the relative entropy. As I show later on, the parameter $\omega \in (0, 1]$ can be interpreted as the share of the cost that is due to standard fixed “menu” costs. Assuming $\bar{F}$ is a continuous probability distribution and letting $\bar{f}$ be its density, the optimization problem in (4.2) can be solved analytically as $F'(x) = f(x) \propto \bar{f}(x) \exp(V(x)/[\theta(1 - \omega)])$, and letting $V^{\max} = \max_x V(x)$:

$$V^* = V^{\max} + \theta(1 - \omega) \log \left(\int_{-\infty}^{\infty} \bar{f}(z) \exp((V(z) - (V^{\max}))/[\theta(1 - \omega)]) dz \right) - \omega \hat{c}(\theta), \quad (\text{OA.B.2})$$

which holds for $\omega < 1$, and just $V^* = V^{\max} - \hat{c}(\theta)$ if $\omega = 1$. The value of adjusting is thus $V^* = V^{\max} - c(\omega, \theta, \bar{f}, V)$ where $c(\omega, \theta, \bar{f}, V)$ is the cost of price adjustment. This resembles the standard menu cost model, with the difference that the cost is endogenous. Letting $\hat{c}(\theta) = \theta |\log(\theta)|/4$, I characterize the solution when $\theta \rightarrow 0$ in the next set of propositions.

Proposition OA.1. *Inaction Regions.* *Assume that $\mathcal{C}(F) = \omega \hat{c}(\theta) + \theta(1 - \omega) \mathcal{I}(F || \bar{F})$ and (V, V^*, F) solve (4.1)-(OA.B.2). Fix $\omega \in (0, 1]$. Then, as $\theta \rightarrow 0$ the total cost of changing prices, $c(\omega, \theta, \bar{f}, V)$ becomes:*

$$c(\omega, \theta, \bar{f}, V) = \hat{c}(\theta) + o(\theta \log(\theta)). \quad (\text{OA.B.3})$$

Furthermore, there is a nontrivial inaction region $(-\bar{x}, \bar{x})$ with:

$$\bar{x} = \left(\frac{6\sigma^2}{B} \times \hat{c}(\theta) \right)^{1/4} (1 + o(1)). \quad (\text{OA.B.4})$$

In particular, $\bar{x}$ does not depend on ω at leading order.

Proof. See Appendix OA.C.1. The key complication of the proof is to show that, despite endogeneity, we obtain (OA.B.3). Then (OA.B.4) follows from the standard argument in Dixit (1991). $\square$

This proposition shows that as $\theta \rightarrow 0$, the total cost converges to $\hat{c}(\theta)$. Since the standard menu cost component is $\omega\hat{c}(\theta)$, ω corresponds to the share of adjustment costs that comes from menu costs. The mixed model is capable of generating inaction as long as there is some fixed menu cost component ($\omega > 0$).³ Furthermore, the size of the inaction region is *independent* of the breakdown between menu and control costs at leading order. As a corollary, inaction regions still scale with the quartic root of costs, so very small costs can generate large inaction regions (Akerlof and Yellen, 1985; Dixit, 1991).

More broadly, as long as there is some menu cost, this result shows that both menu and control costs are “perfect substitutes” for small values of θ : regardless of the breakdown, they generate exactly the same inaction region at leading order. This result directly relates to the sources of price changing costs. While the name “menu-costs” refers to literal physical costs of changing prices, it is widely believed that these costs actually reflect managerial time or information processing costs.⁴ The result shows that the distinction is immaterial in terms of the size of the inaction bands.

As the next proposition shows, however, there is a key difference between both costs: as $\omega < 1$, there is a nontrivial asymptotic distribution for the reset gaps. The proposition also shows that once properly scaled, reset gaps are asymptotically normal.

Proposition OA.2. *Reset Gap Distribution and Expected Value upon Reset.* *As $\theta \rightarrow 0$,*

$$|\theta/\log(\theta)|^{-1/4}x^* \rightarrow^d \mathcal{N}(0, (1-\omega)\sqrt{3\sigma^2/2B}).$$

The expected value upon resetting without accounting for the costs, $\mathbb{E}[V] = \int_{-\infty}^{\infty} f(x)V(x)dx$ is

$$\mathbb{E}[V] = V^{\max} - \frac{(1-\omega)\theta}{2} + o(\theta) = V^{\max} + o(\theta|\log(\theta)|).$$

Proof. See Appendix OA.C.1. □

We can directly see from the formula how different structural parameters map into the size of mistakes. First, a higher share of control costs $(1-\omega)$ increases the asymptotic variance: if precision is more costly, firms choose less precisely. If B is larger, the variance of mistakes will be smaller: if the profit function has higher curvature, mistakes are more

³If $\omega = 0$, firms can sidestep adjustment costs by sampling infinitely often from the baseline distribution for free. Once $\omega > 0$, this is no longer optimal and a standard inaction region emerges. This will not be an issue in a discrete time setup, although in that case the definition of the time period is important, see the discussion in Costain and Nakov (2015).

⁴For example, Stokey (2009) says that the fixed cost of changing prices is “*interpreted as the time cost of managerial decision making as well as the cost of printing a new menu*”.

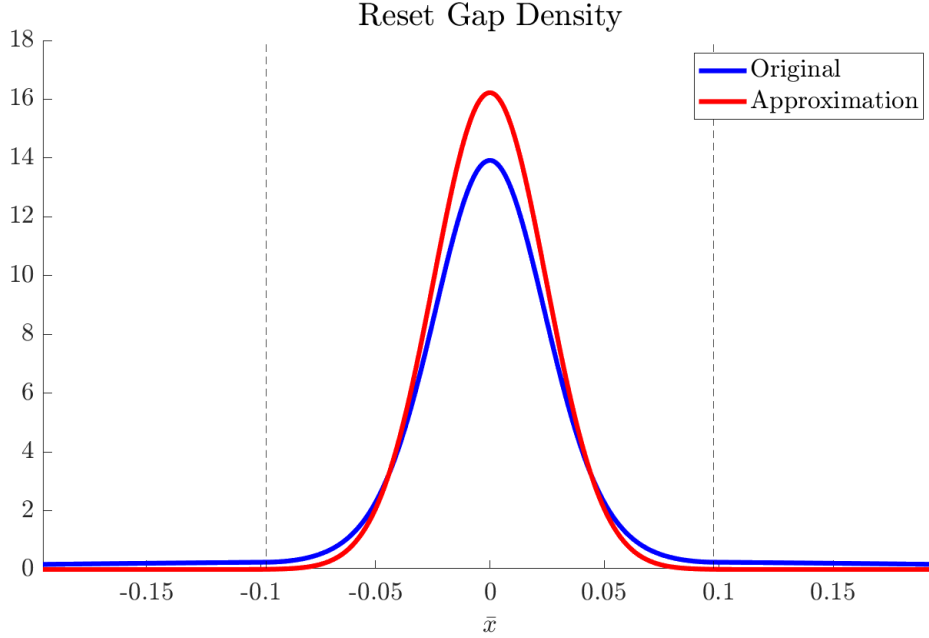


Figure OA.B.1: Reset gap distribution, global solution (blue) and analytical approximation (red).

costly, and thus the firm will optimally choose to minimize them by paying higher control costs. Conversely, if σ^2 is larger, the firm will also make larger mistakes. Intuitively, larger σ^2 means that prices will change more often, and thus the firm optimally economizes on price adjustment costs by having more dispersed reset prices.

Note that, since the convergence rate is quite slow, relatively large mistakes in price space can arise even for very small information costs (as indexed by θ). This is conceptually similar to Akerlof and Yellen (1985): since the value function is locally flat close to 0, deviations from the optimal price that are “large” in the policy space lead to “small” losses in profits. This provides a theoretical argument for why relatively large ex-post imperfections could arise even if control costs were small in practice.

OA.B.4 Quality of the Approximation

In this section I quantify the quality of the approximation. I set $\sigma^2 = (0.1)^2$, $\rho = 0.04$, and $\omega = 0.5$, so half the cost is the typical menu cost and the other half is the control cost. To compute the global solution, I need to specify a baseline distribution $\bar{f}(x)$. I use a normal distribution with mean zero and standard deviation 0.2, which is quite large relative to the other parameters.

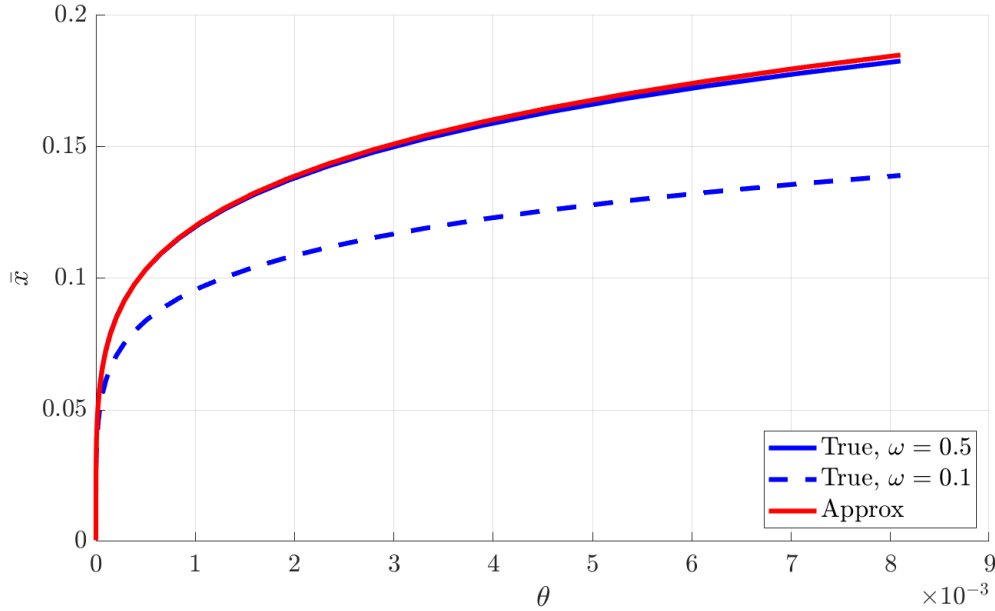


Figure OA.B.2: Global (blue) and approximate (red) inaction thresholds as a function of θ . Solid: $\omega = 0.5$, dashed: $\omega = 0.1$.

Figure OA.B.1 shows the reset gap distribution for $\theta = (0.14)^4$, which generates inaction bands of around 9.8 percent and a frequency of price changes of $N = 1.2$, quite close to the UK data (where the average size is 9.8% and the frequency $N = 1.3$). As we can see, the approximation is good even at these sizes. The true distribution has higher variance, and in particular has higher probability of the reset gap falling outside of the inaction region. The intuition for this is that, given that getting a more precise draw is costly, a firm may find it “cheaper” to have many less precise draws (and, if we fall outside of the inaction region, we reset again) than one very precise draw. Note also that, quantitatively, there is some decent dispersion in reset prices.

Figure OA.B.2 shows the accuracy of the approximation for different values of ω . As we can see, when $\omega = 0.5$, the approximation is quite good for a large range of $\bar{x}$. However, once we consider smaller ω , the approximation quickly overestimates the true inaction region. The intuition is related to the phenomenon discussed above: if ω is small, then the fixed cost of each draw is small, so the firm finds it optimal to pay this fixed cost several times and sample from relatively bad distributions, rather than to pay the fixed cost once and sample from a very precise distribution. Thus, the same phenomenon that makes basic menu costs necessary to generate inaction also appears (in a less extreme form) when those fixed costs

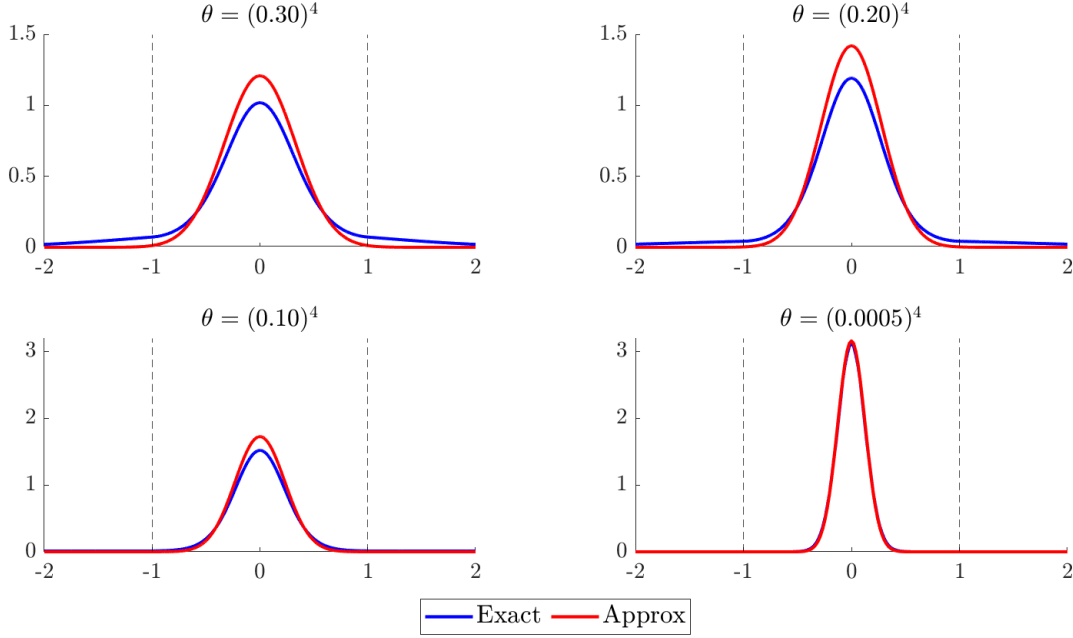


Figure OA.B.3: Global (blue) and approximate (red) reset distribution for different values of θ . Horizontal axis normalized by the size of the inaction bands, i.e., 2 means $x = 2\bar{x}$.

are small. This shows that the necessity of some kind of additional force for inaction is not a bug of continuous time, but rather a characteristic of this type of model, in which price resets come with randomness but we immediately realize we made an error, and thus we can potentially minimize pricing costs by sampling repeatedly.

Finally, Figure OA.B.3 shows how the true distribution converges to the analytical approximation as θ becomes smaller. The x axis of all plots is in terms of the size of the inaction band, i.e., 1 means $x = \bar{x}$, 2 means $x = 2\bar{x}$ and so on. As we can see, the distribution becomes more concentrated at 0 as θ becomes smaller. This follows from our earlier discussion on speeds of convergence: bands converge more slowly than the asymptotic distribution, so even if we scale x by $\bar{x}$ we will still see the distribution converging to a point mass at 0. This is also why the uncertainty about the reset price does not add an additional cost that affects the inaction band: at this order, uncertainty goes to zero faster than the other costs and thus does not matter.

OA.B.5 Extended Identification Result

This section provides an extended identification result. Under some regularity assumptions, it shows how to fully recover $F(x)$ from appropriate generalizations of the regressions I run

in the main text. Furthermore, in the case of $n = 1$, I provide an extension of the results in Alvarez et al. (2021) that allows me to recover the hazard $\Lambda(x)$ from the price-change distribution $q(\Delta p)$. Several of the proofs use similar techniques to that of Proposition 10, which is covered in Appendix C.7.

OA.B.5.1 Identification of the Reset Gap Distribution

Consider the environment of the main text. Write $\Delta\tau_{i,n} = \tau_{i,n} - \tau_{i,n-1}$ for the duration between consecutive adjustments, and interpret $\mathcal{F}_{\tau_{i,n-1}}$ as the information immediately after the previous reset, including $x_{i,n-1}^*$ and $\Delta p_{i,n-1}$. Assume throughout that the moments and weighted expectations used below are finite and that conditional optional stopping is valid. Then:

$$\begin{aligned}\mathbb{E}[(\Delta p_{i,n})^2 - \sigma^2 \Delta\tau_{i,n} | \mathcal{F}_{\tau_{i,n-1}}] &= \mathbb{E}[(x_{i,n}^* - x_{i,n-1}^*)^2 | \mathcal{F}_{\tau_{i,n-1}}] \\ &\quad - 2\sigma \mathbb{E}[(x_{i,n}^* - x_{i,n-1}^*)(W_{i,\tau_{i,n}} - W_{i,\tau_{i,n-1}}) | \mathcal{F}_{\tau_{i,n-1}}] \\ &\quad + \sigma^2 \mathbb{E}[(W_{i,\tau_{i,n}} - W_{i,\tau_{i,n-1}})^2 - \Delta\tau_{i,n} | \mathcal{F}_{\tau_{i,n-1}}] \\ &= \mathbb{E}[(x^*)^2] + (x_{i,n-1}^*)^2\end{aligned}$$

The final expectation is zero by conditional optional stopping for the squared Brownian increment minus elapsed time, exactly as in Appendix C.7. The middle expectation is zero because the stopped Brownian increment has zero conditional mean and the reset $x_{i,n}^*$ is independent of the completed spell and has mean zero. Note that the conditional expected duration can depend on the initial reset gap. Consider now a regression of this corrected outcome on $(\Delta p_{i,n-1})^{2k}$ for $k \geq 1$, with an intercept. Multiplying its slope by the variance of the regressor recovers the covariance:

$$\begin{aligned}\text{Cov}((\Delta p_{i,n-1})^{2k}, (\Delta p_{i,n})^2 - \sigma^2 \Delta\tau_{i,n}) &= \mathbb{E}[(\Delta p_{i,n-1})^{2k} ((\Delta p_{i,n})^2 - \sigma^2 \Delta\tau_{i,n})] \\ &\quad - \mathbb{E}[(\Delta p_{i,n-1})^{2k}] \mathbb{E}[(\Delta p_{i,n})^2 - \sigma^2 \Delta\tau_{i,n}].\end{aligned}$$

By stationarity and symmetry, we can directly estimate the second term as:

$$\mathbb{E}[(\Delta p_{i,n})^{2k}] \mathbb{E}[(\Delta p_{i,n})^2 - \sigma^2 (\tau_{i,n} - \tau_{i,n-1})] = \mathbb{E}[(\Delta p)^{2k}] (\text{Var}(\Delta p) - \sigma^2 / N).$$

I now focus on the first term. Note that using the law of iterated expectations and conditioning on $\mathcal{F}_{\tau_{i,n-1}}$:

$$\begin{aligned}\mathbb{E}[(\Delta p_{i,n-1})^{2k}((\Delta p_{i,n})^2 - \sigma^2(\tau_{i,n} - \tau_{i,n-1}))] &= \mathbb{E}[(\Delta p_{i,n-1})^{2k}(\mathbb{E}[(x^*)^2] + (x_{i,n-1}^*)^2)] \\ &= \text{Var}(x^*)\mathbb{E}[(\Delta p)^{2k}] + \mathbb{E}[(\Delta p_{i,n-1})^{2k}(x_{i,n-1}^*)^2]\end{aligned}$$

but since

$$\begin{aligned}\mathbb{E}[(\Delta p_{i,n-1})^{2k}(x_{i,n-1}^*)^2] &= \mathbb{E}[(x_{i,n-1}^* - x_{i,\tau_{i,n-1}})^{2k}(x_{i,n-1}^*)^2] \\ &= \mathbb{E}\left[\left(\sum_{j=0}^{2k} \binom{2k}{j} (x_{i,n-1}^*)^{2k-j} (-x_{i,\tau_{i,n-1}})^j\right) (x_{i,n-1}^*)^2\right]\end{aligned}$$

we can use the fact that $x_{i,n-1}^*$ is independent of $-x_{i,\tau_{i,n-1}}$ to write this as:

$$\mathbb{E}[(\Delta p_{i,n-1})^{2k}(x_{i,n-1}^*)^2] = \sum_{j=0}^{2k} \binom{2k}{j} \mathbb{E}[(x_{i,n-1}^*)^{2(k+1)-j}] \mathbb{E}[-x_{i,\tau_{i,n-1}}]^j$$

now, by symmetry, the odd values of j just yield zero. Thus we can write:

$$\mathbb{E}[(\Delta p_{i,n-1})^{2k}(x_{i,n-1}^*)^2] = \sum_{\ell=0}^k \binom{2k}{2\ell} \mathbb{E}[(x_{i,n-1}^*)^{2(k+1-\ell)}] \mathbb{E}[(x_{i,\tau_{i,n-1}})^{2\ell}].$$

The minus sign disappears because only even powers remain. Thus, the covariances identify sums of products of moments of x^* and $x_{i,\tau_{i,n-1}}$. I show below how to use these expressions to recover those moments separately using appropriate regression coefficients.

OBSERVABLE REGRESSION COEFFICIENTS. For each $k \geq 1$, consider the two separate population OLS regressions, each including an intercept:

$$(\Delta p_{i,n})^2 = \alpha_{2,k} + \tilde{\beta}_{2,k}(\Delta p_{i,n-1})^{2k} + \text{Error}, \quad (\text{OA.B.5})$$

$$\Delta \tau_{i,n} = \alpha_{3,k} + \tilde{\beta}_{3,k}(\Delta p_{i,n-1})^{2k} + \text{Error}. \quad (\text{OA.B.6})$$

At $k = 1$, these are the regressions in the main text: $\tilde{\beta}_{2,1} = \tilde{\beta}_2$ and $\tilde{\beta}_{3,1} = \tilde{\beta}_3$. Define the observable corrected covariance

$$\begin{aligned} C_k &= (\tilde{\beta}_{2,k} - \sigma^2 \tilde{\beta}_{3,k}) \text{Var}((\Delta p)^{2k}) \\ &= \text{Cov}((\Delta p_{i,n-1})^{2k}, (\Delta p_{i,n})^2 - \sigma^2 \Delta \tau_{i,n}). \end{aligned}$$

Thus the slopes are used together with the marginal price-change moments. In particular, $\text{Var}((\Delta p)^{2k}) = \mathbb{E}[(\Delta p)^{4k}] - \mathbb{E}[(\Delta p)^{2k}]^2$. If this variance is zero, the slope is undefined but the covariance itself is zero and can be used directly. The duration correction retains the covariance between duration and the previous price change; it does not set it to zero.

RESET AND PRE-RESET MOMENTS. In order to save notation, let

$$\mu_{2k} = \mathbb{E}[(x^*)^{2k}], \quad \nu_{2k} = \mathbb{E}[(x_{i,\tau_{i,n-1}})^{2k}], \quad \mu_0 = \nu_0 = 1.$$

be the corresponding moments. A similar argument to above (expanding the square and using independence) yields:

$$\begin{aligned} \mathbb{E}[(\Delta p)^{2k}] &= \mathbb{E}[(x^* - x_{i,\tau_i})^{2k}] \\ &= \sum_{\ell=0}^k \binom{2k}{2\ell} \mu_{2(k-\ell)} \nu_{2\ell}. \end{aligned}$$

Using that $\mathbb{E}[(\Delta p)^2 - \sigma^2 \Delta \tau] = 2\mu_2$ and replacing in the covariance above yields

$$C_k = \sum_{\ell=0}^k \binom{2k}{2\ell} \mu_{2(k+1-\ell)} \nu_{2\ell} - \mu_2 \mathbb{E}[(\Delta p)^{2k}].$$

The coefficient on the new reset moment μ_{2k+2} is one. All the other terms can be computed recursively.

ITERATIVE PROCEDURE. The linear predictability regression in the main text identifies

$$\mu_2 = -\beta_1 \text{Var}(\Delta p), \quad \sigma^2 = N \text{Var}(\Delta p)(1 + 2\beta_1), \quad N = \frac{1}{\mathbb{E}[\Delta \tau]}.$$

These quantities are known before constructing any C_k . Starting from $\mu_0 = \nu_0 = 1$ and the identified μ_2 we can identify all even moments of x^* and $x_{i,\tau_{i,n-1}}$ by iterating following two steps for $k = 1, 2, \dots$:

$$\nu_{2k} = \mathbb{E}[(\Delta p)^{2k}] - \mu_{2k} - \sum_{\ell=1}^{k-1} \binom{2k}{2\ell} \mu_{2(k-\ell)} \nu_{2\ell}, \quad (\text{OA.B.7})$$

$$\mu_{2k+2} = C_k + \mu_2 \mathbb{E}[(\Delta p)^{2k}] - \sum_{\ell=1}^k \binom{2k}{2\ell} \mu_{2(k+1-\ell)} \nu_{2\ell}, \quad (\text{OA.B.8})$$

where I follow the convention that empty sums are zero. At the start of step k , moments of x^* through order $2k$ and moments of $x_{i,\tau_{i,n-1}}$ through order $2k-2$ have already been recovered. The first equation computes ν_{2k} using the observed marginal moment and those previously recovered moments. The second equation then computes μ_{2k+2} using the corrected regression covariance and the newly recovered pre-reset moment. Neither step requires knowing the hazard or either latent gap distribution in advance. For example, the first iteration gives

$$\begin{aligned} \nu_2 &= \mathbb{E}[(\Delta p)^2] - \mu_2, \\ \mu_4 &= C_1 + \mu_2^2 = \mu_2^2 + (\tilde{\beta}_2 - \sigma^2 \tilde{\beta}_3) \text{Var}((\Delta p)^2), \end{aligned}$$

which reproduces the fourth-moment result in the main text. The next iteration gives

$$\begin{aligned} \nu_4 &= \mathbb{E}[(\Delta p)^4] - \mu_4 - 6\mu_2\nu_2, \\ \mu_6 &= C_2 + \mu_2 \mathbb{E}[(\Delta p)^4] - 6\mu_4\nu_2 - \mu_2\nu_4. \end{aligned}$$

Thus the regressions with the fourth power of the lagged price change identify the sixth reset moment, after the second and fourth moments have been recovered.

UNIQUENESS OF THE RESET DISTRIBUTION. The recursion identifies every even reset moment from the even moments of the unconditional price change distribution, mean duration, the linear predictability coefficient, and the sequence of squared-change and duration regression coefficients. Since all odd moments are zero due to symmetry, the above construction allows identification of all moments of F . It is not generally true that identifying all of the moments automatically identifies a distribution. Thus, we need to further restrict the admissible F distributions such that the moments uniquely identify F .

Assumption OA.1. *The distribution F has finite moments of every order and is uniquely*

determined by its moments.

One sufficient condition that guarantees this is Carleman's condition, which states that if the moments satisfy $\sum_{k=1}^{\infty} \mu_{2k}^{-1/2k} = +\infty$, then the distribution can be uniquely recovered from them. This condition essentially restricts very fat-tailed distributions. The generalized normal distribution I use in Section 5.6 satisfies this condition whenever $d \geq 1$, which is what I find empirically.

OA.B.5.2 Identifying the Hazard given the Reset Gap Distribution

The above result shows how to point identify F by using regressions (OA.B.5)-(OA.B.6) and appropriate unconditional moments. In this subsection, I show how to recover $\Lambda(x)$ by using the identified F in conjunction with the distribution of price changes $q(\Delta p)$. Throughout the section, I assume that $n = 1$, as in Alvarez et al. (2021).

Assume that F is absolutely continuous with density f , and assume that $\bar{x} = +\infty$, so there is no adjustment barrier. Assume also that $\sigma^2 > 0$ and $0 < N < \infty$. For the Fourier calculations below, assume that g is twice continuously differentiable, $g, g', g'' \in L^1(\mathbb{R})$, and $g(x), g'(x) \rightarrow 0$ as $|x| \rightarrow \infty$. Finally, assume $g(x) > 0$ for all x , so that the division by the stationary density is well-defined. The distribution of price changes is $q(\Delta p) = \int \frac{\Lambda(z-\Delta p)g(z-\Delta p)}{N} f(z)dz = \int \frac{\Lambda(x)g(x)}{N} f(x + \Delta p)dx$ since $\Delta p = x^* - x$, and thus for each possible Δp we integrate across pairs of x and x^* that could have generated it. Let $\hat{h}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} h(x)dx$ be the Fourier transform of a function h . Fourier transforms are useful here because of two properties: i) the Fourier transform of a convolution equals the product of the Fourier transform, and ii) they operate conveniently with derivatives, in the sense that the Fourier transform of $h'(x)$ is $i\omega \hat{h}(\omega)$. Using these properties, we can obtain that:

$$N\hat{q}(\omega) = \hat{f}(\omega)[\hat{\Lambda}g](\omega),$$

so knowledge of $\hat{f}(\omega)$ and the distribution of price changes $\hat{q}(\omega)$ allows us to obtain $[\hat{\Lambda}g](\omega)$ and thus $[\Lambda g](x)$ (i.e their product at each point) via the inverse Fourier transform. The quotient is properly defined, since the moment-determinacy condition we have implies that

$\hat{f}(\omega) \neq 0$.^{5,6} In order to identify $\Lambda(x)$, we need to separate g from Λ . Using the KFE, $\Lambda(x)g(x) = \frac{\sigma^2}{2}g''(x) + Nf(x)$ so:

$$Nq(x) = \int \left[\frac{\sigma^2}{2}g''(z) + Nf(z) \right] f(x+z) dz.$$

Using the property for the derivative of Fourier transforms, we get:

$$N\hat{q}(\omega) = N\hat{f}(\omega)^2 - \frac{\sigma^2}{2}\omega^2\hat{f}(\omega)\hat{g}(\omega)$$

where by symmetry of f and g the complex conjugate equals the function. Since we know f , we know $\hat{f}(\omega)$, and thus we can recover the stationary density of gaps $\hat{g}(\omega)$ as:

$$\hat{g}(\omega) = \frac{2N}{\sigma^2} \frac{\hat{f}(\omega)^2 - \hat{q}(\omega)}{\omega^2 \hat{f}(\omega)}$$

for $\omega \neq 0$ with $\hat{f}(\omega) \neq 0$. At zero, the second-moment identity implies

$$\hat{f}(\omega)^2 - \hat{q}(\omega) = \frac{\mathbb{E}[(\Delta p)^2] - 2\mu_2}{2}\omega^2 + o(\omega^2) = \frac{\sigma^2}{2N}\omega^2 + o(\omega^2).$$

Since $\hat{f}(0) = 1$, the quotient converges to one. Thus we set $\hat{g}(0) = 1$ and recover its values at any other zeros of $\hat{f}$ by continuity, as above. Once we have $\hat{g}(\omega)$, we can use the inverse Fourier transform to recover $g(x)$, and then divide $[\Lambda g](x)$ over $g(x)$ to recover $\Lambda(x)$ almost everywhere.

OA.B.6 Additional Discussion on the Structural Model

Here I present some additional discussion on the structural models in Section 5.6.

⁵If $\hat{f}$ was equal to zero on an open interval, all its derivatives would equal zero at an interior point ω_0 . Then $f(x)[1 + \varepsilon \cos(\omega_0 x)]$, for $0 < |\varepsilon| < 1$, would be a different symmetric density with exactly the same moments as f , contradicting moment determinacy. Thus no additional nonvanishing assumption on $\hat{f}$ is needed.

⁶Alternatively, the construction done in the previous part also gives the moments of the the distribution of $x_{i,\tau_{i,n}-1}$. This distribution has a density proportional to $[\Lambda g](x)$. However, in order to use those moments to point identify it, we would need to use similar assumptions of moment determinacy, in this case for the density $\Lambda g/N$, which is an endogenous object.

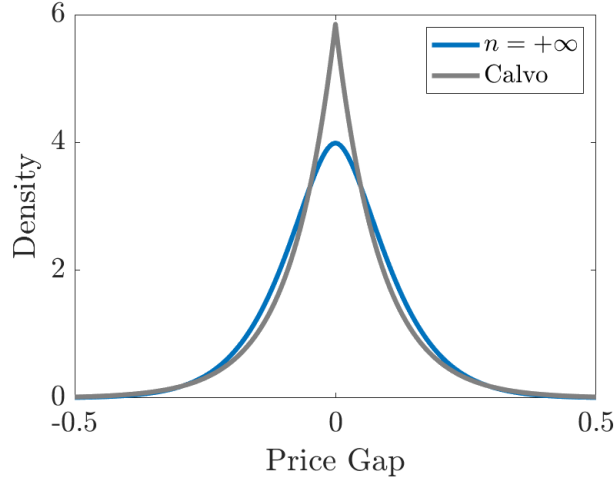


Figure OA.B.4: Stationary Distribution of Price Gaps. Blue: model with $n \rightarrow +\infty$. Grey: Standard Calvo model (no mistakes), recalibrated to match the same variance of price changes.

ON THE KURTOSIS OF A CALVO MODEL WITH IMPERFECT RESET. Some readers may find the model's kurtosis below six surprising, since it is well known that, in models with perfect reset, Calvo hazards imply a kurtosis of 6, the maximum possible in this class of models (Alvarez et al., 2016a). There are two sources for this discrepancy: first, variation in realized price changes is driven by both variation in the price gap before reset and in the reset point. That directly impacts the measured kurtosis, as shown in Lemma 3. Second, the distribution of price gaps before reset is also affected by the randomness in reset gaps. Figure OA.B.4 shows the stationary distribution of price gaps (which with constant hazards equals the distribution of price gaps conditional on an ended spell) for the baseline model as well as for a standard Calvo model with deterministic reset point at $x = 0$, recalibrated to match the same $\text{Var}(\Delta p)$. The figure shows that the randomness in the reset point smooths out the distribution at 0. This pushes mass away from zero towards intermediate values of x , thus lowering the kurtosis of the distribution of price gaps at reset, which further lowers the kurtosis of the distribution of price changes.

FIT OF A CALVO MODEL ON THE EXTENDED SET OF MOMENTS. While our baseline Calvo model fits N , $\text{Kurt}(\Delta p)$, β_1 and β_2 quite well, the reader may wonder how this model fits $\tilde{\beta}_2$ and $\tilde{\beta}_3$ separately. First, note that as $n \rightarrow +\infty$, the timing of the price change i does not depend on the individual $x_{i,t}$ of that product, and thus $\tilde{\beta}_3 = 0$ for any parametrization. Empirically, the specifications in Appendix A paint a consistent picture: $\tilde{\beta}_3$ is negative and

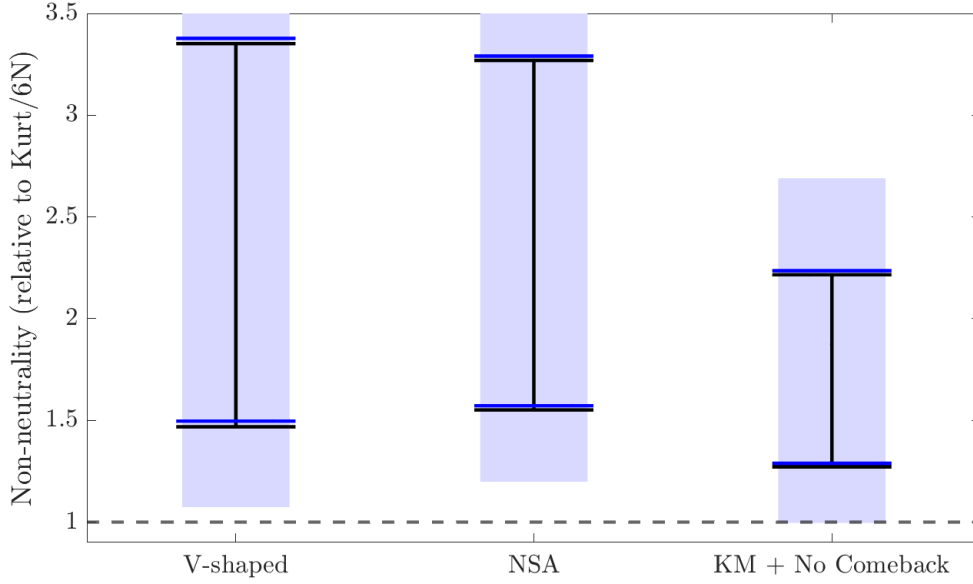


Figure OA.B.5: Estimated bounds relative to $\text{Kurt}(\Delta p)/6N$ (black), and estimates if we set $\tilde{\beta}_3 = 0$ in each realization (blue).

statistically significant. At the baseline specification, we have $\hat{\tilde{\beta}}_3 = -0.178$.

However, while statistically significant, its estimated values are economically very small. In order to make that point clear, Figure OA.B.5 recomputes the bounds as in Figure 5.1a, but setting $\tilde{\beta}_3 = 0$ at each bootstrap draw. The resulting bounds are shown in blue. As we can see, the change is minimal.

I also complement this evidence with some reduced-form simulations. In particular, I consider models with $n = 1$, with a hazard $\Lambda(x) = \lambda_0(1 + \lambda_1 x^2)$ where different values of λ_1 captures different degrees of state dependence. I assume that $\sigma^2 = 0.0082$, and that F is a generalized normal with $b = 0.0726$ and $d = 1.11$. These are approximately the point estimates underlying the baseline model of Section 5.6. I consider four alternative parametrizations for the hazard: 1) a Calvo parametrization, with $\lambda_0 = 1.3$ and $\lambda_1 = 0$; 2) a parametrization with low state-dependence, where $\lambda_0 = 1.28$ and $\lambda_1 = 0.3/(0.4^2)$; 3) a parametrization with state-dependence calibrated to roughly match the estimated hazard in Gagliardone et al. (2025), which yields $\lambda_0 = 1.17$ and $\lambda_1 = 1.5/(0.4^2)$; and 4) a calibration with high state-dependence with $\lambda_0 = 0.68$ and $\lambda_1 = 20/(0.4^2)$. The values of λ_0 are chosen such that all calibrations match a similar average frequency of price changes. I simulate the price paths of 50000 firms for fifty years, where each time interval is set to a day ($dt = 1/360$).

I burn the first ten years of the simulation, and then estimate the moments of interest on simulated data.

	Calvo	Low SD	GG	High SD
$\tilde{\beta}_3$ (simulation)	0.009844	-0.229456	-0.771790	-2.649186

Table OA.4: Simulated moments across calibrations. “Low SD” has $\lambda_1 = 0.3/(0.4^2)$, “GG” (Gagliardone-Gertler) has $\lambda_1 = 1.5/(0.4^2)$, “High SD” has $\lambda_1 = 20/(0.4^2)$. All calibrations use λ_0 adjusted to target frequency ≈ 1.3 .

Table OA.4 contains the results. As we can see, the low state-dependence calibration roughly matches the estimated magnitude of $\tilde{\beta}_3$. Under this calibration, a price that is off by 40 percentage points is only 30% more likely to adjust than a price that is set at $x = 0$. Another way to see that state-dependence is very low is that, under this calibration, $\Lambda(0) = 1.28$, while the average frequency of price changes is around 1.3. Thus, this features essentially Calvo adjustment. The table also shows how other calibrations that feature higher state-dependence would imply much larger magnitudes for $\tilde{\beta}_3$, which are counterfactual. Overall, this shows that, while statistically significant, the magnitude of $\tilde{\beta}_3$ is small, which is consistent with models with very low state-dependence.

OA.B.7 Examples for Proposition 9

In all cases, take $n = 1$ so we are in a single-product case.

$CIR = +\infty$. Assume:

$$\Lambda(x) = \begin{cases} 0 & \text{if } |x| < \bar{x} \\ 2N & \text{if } |x| \geq \bar{x} \end{cases}.$$

Assume $f(x)$ is 0 for all $x \in [-(\bar{x} + \varepsilon_2), \bar{x} + \varepsilon_2]$, and has some distribution (which we later specify) in the rest of the domain. Assume $\sigma^2 = 0$, so firms do not travel along the state space. Given this, there are many invariant distributions (since we do not have mixing). Assume the following symmetric invariant distribution:

$$g(x) = \frac{1}{2}\delta_0 + \frac{1}{2}f(x).$$

That is, half of the firms sit at $x = 0$, and the other half is distributed according to $f(x)$, which only has density in parts where $\Lambda(x) > 0$. Note that this is a proper invariant distribution: for all firms with $|x| < \bar{x}$, they just stay where they are. For firms with $|x| \geq \bar{x}$,

they are always outside the bands, they move with hazard $2N$, and when they do they randomly land in another place with density $f(x)$. The KFE in this outside region reads:

$$\begin{aligned} 2Ng &= Nf \\ Nf(x) &= Nf(x), \end{aligned}$$

so it is trivially satisfied for any N . The price change distribution is given by the distribution of $x_1^* - x_2^*$, so we can match any $\text{Kurt}(\Delta p) > 2$ by choosing $\text{Kurt}(x^*)$. I now show that this setting implies unboundedly high CIR. Assume there is a marginal shock of size δ . Initially all firms shift down in x . Then:

- Firms with $|x| > \bar{x} + \varepsilon_2$ are all in the region where they can change prices, so over time they reset with hazard $2N$, and eventually all of them adjust. Thus the contribution of this term to the CIR is finite.
- Firms with $|x| < \bar{x}$ are displaced, but they never leave the absorbing region. Thus, they never change prices, and have permanently negative x . Thus, $-\mathbb{E}[\int_0^\infty x_s ds | x(0) = x] = +\infty$ for these firms. Since there is a positive mass of such firms, the CIR is $+\infty$.

The key content of the example is that, once we allow errors, we can match pricing moments without needing σ^2 . That implies that some firms can take an unboundedly long time to change prices for the first time, which makes the CIR diverge.

$CIR = -\infty$. Assume again that $\sigma^2 = 0$. Assume that the reset price distribution is of the following form:

$$x^* = \begin{cases} -a & \text{prob } \frac{\varepsilon}{2} \\ -\bar{x} & \text{prob } \frac{1-\varepsilon}{2} \\ \bar{x} & \text{prob } \frac{1-\varepsilon}{2} \\ a & \text{prob } \frac{\varepsilon}{2} \end{cases},$$

with $a > \bar{x}$, both arbitrary for now. For simplicity, I will assume that if a new draw falls into the same spot, it counts as a price change of size 0. Assume the GHF is:

$$\Lambda(x) = \begin{cases} \lambda_1 & \text{if } |x| \leq \bar{x} \\ \lambda_2 & \text{if } |x| > \bar{x} \end{cases},$$

let $\lambda_1 = N \frac{1-\varepsilon}{1-p}$ and $\lambda_2 = N \frac{\varepsilon}{p}$ where p is also chosen by us. Under these assumptions, the invariant distribution is given by:

$$\check{x} = \begin{cases} -a & \text{prob } \frac{p}{2} \\ -\bar{x} & \text{prob } \frac{1-p}{2} \\ \bar{x} & \text{prob } \frac{1-p}{2} \\ a & \text{prob } \frac{p}{2} \end{cases}.$$

The key thing to notice is that since $\sigma^2 = 0$, there are only 4 values where we can be: $\pm a, \pm \bar{x}$. In steady state, the mass leaving each point is $\text{Prob}(\check{x} = y)\Lambda(y)$, and the mass entering is $N\text{Prob}(x^* = y)$. Given our choice of the hazard function, this always holds, so the mentioned distribution is invariant. It is straightforward to show that under this hazard, the frequency is exactly N . Given that, the second step is to compute the $m(x)$ function. Given that x is fixed and the reset probability is known, we can compute it directly as:

$$\begin{aligned} m(x) &= -\mathbb{E}\left[\int_0^\infty x(t)dt | x(0) = x\right] \\ &= -x \int_0^\infty e^{-\lambda_i} dt = -\frac{x}{\lambda_i}. \end{aligned}$$

Given that, we can compute the CIR as:

$$\mathcal{M}(\delta) = \int_{-\infty}^\infty m(x - \delta) dG_{\check{x}},$$

for any $\delta > 0$, direct computation shows:

$$\begin{aligned} \mathcal{M}(\delta) &= -\frac{p}{2}\left(\frac{1}{\lambda_2}(-a - \delta) + \frac{1}{\lambda_2}(a - \delta)\right) - \frac{1-p}{2}\left(\frac{1}{\lambda_2}(-\bar{x} - \delta) + \frac{1}{\lambda_1}(\bar{x} - \delta)\right) \\ &= \delta\left(\frac{p}{\lambda_2} + \frac{1-p}{2}\left(\frac{1}{\lambda_2} + \frac{1}{\lambda_1}\right)\right) - \bar{x}\frac{1-p}{2}\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right), \end{aligned}$$

whereas for any $\delta < 0$, the same computation shows:

$$\begin{aligned} \mathcal{M}(\delta) &= -\frac{p}{2}\left(\frac{1}{\lambda_2}(-a - \delta) + \frac{1}{\lambda_2}(a - \delta)\right) - \frac{1-p}{2}\left(\frac{1}{\lambda_1}(-\bar{x} - \delta) + \frac{1}{\lambda_2}(\bar{x} - \delta)\right) \\ &= \delta\left(\frac{p}{\lambda_2} + \frac{1-p}{2}\left(\frac{1}{\lambda_2} + \frac{1}{\lambda_1}\right)\right) + \bar{x}\frac{1-p}{2}\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right), \end{aligned}$$

where $\bar{x}^{\frac{1-p}{2}}(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}) > 0$, since $\lambda_2 > \lambda_1$. On the other hand, at exactly $\delta = 0$ we have $\mathcal{M}(0) = 0$ by construction. Thus, in the limit:

$$\lim_{\delta \rightarrow 0^+} \frac{\mathcal{M}(\delta) - \mathcal{M}(0)}{\delta} = \frac{p}{\lambda_2} + \frac{1-p}{2}(\frac{1}{\lambda_2} + \frac{1}{\lambda_1}) - \frac{\bar{x}^{\frac{1-p}{2}}(\frac{1}{\lambda_1} - \frac{1}{\lambda_2})}{\lim_{\delta \rightarrow 0^+} \delta} = -\infty,$$

and

$$\lim_{\delta \rightarrow 0^-} \frac{\mathcal{M}(\delta) - \mathcal{M}(0)}{\delta} = \frac{p}{\lambda_2} + \frac{1-p}{2}(\frac{1}{\lambda_2} + \frac{1}{\lambda_1}) + \frac{\bar{x}^{\frac{1-p}{2}}(\frac{1}{\lambda_1} - \frac{1}{\lambda_2})}{\lim_{\delta \rightarrow 0^-} \delta} = -\infty,$$

thus, the marginal CIR is minus infinity. I show at the end of this section that any kurtosis of price changes above 2 can be matched by appropriately choosing ε and a . The idea is that, in this case, there is a mass of firms very close to an inaction threshold. Since $\sigma^2 = 0$, in steady state firms move in and out, but are always right at the threshold: there is no diffusion term, so they just jump somewhere else whenever they get hit by the (generalized) Calvo fairy. Once we do a marginal shock, these firms cross the inaction threshold, and thus are much more likely to change prices. Since this mass of firms changes prices by a non-zero amount, $\mathcal{M}(\delta)$ jumps at 0: any small monetary shock causes a positive mass of firms to change prices. For a positive shock, this implies that a marginal monetary shock generates a *jump* up in the price level, which implies a jump *down* in output. This makes the marginal CIR “ $-\infty$ ”.

MATCHING ANY KURTOSIS ABOVE 2. In order to show this, note that $\Delta p = x_2^* - x_1^*$ so $\text{Kurt}(\Delta p) = \frac{1}{2}(\text{Kurt}(x^*) + 3)$. Thus, if we can generate any desired value (above 1) for $\text{Kurt}(x^*)$, we are done. Say we want a kurtosis $\text{Kurt}(x^*) = K$. Then direct computation with this distribution shows:

$$K = \frac{a^4 \varepsilon + \bar{x}^4 (1 - \varepsilon)}{(a^2 \varepsilon + \bar{x}^2 (1 - \varepsilon))^2}.$$

Letting $z = a^2$, for any $K \geq 1$ this describes a quadratic:

$$K(\varepsilon z + \bar{x}^2(1 - \varepsilon))^2 = z^2 \varepsilon + \bar{x}^4(1 - \varepsilon),$$

if this has any positive real roots, we are done. Operating:

$$K(\varepsilon^2 z^2 + 2\bar{x}^2(1 - \varepsilon)\varepsilon z + \bar{x}^4(1 - \varepsilon)^2) = z^2 \varepsilon + \bar{x}^4(1 - \varepsilon)$$

$$\begin{aligned}
z^2(\varepsilon(K\varepsilon - 1)) + zK(2\bar{x}^2(1 - \varepsilon)\varepsilon) + \bar{x}^4(1 - \varepsilon)(K(1 - \varepsilon) - 1) &= 0 \\
zK(2\bar{x}^2(1 - \varepsilon)\varepsilon) + \bar{x}^4(1 - \varepsilon)(K(1 - \varepsilon) - 1) &= z^2\varepsilon(1 - K\varepsilon).
\end{aligned}$$

since $\varepsilon > 0$, the left-hand side is an upward-sloping line. The RHS is a parabola centered at 0, which is U-shaped whenever $1 - K\varepsilon > 0$. Since we can make ε arbitrarily small, just take $\varepsilon = b\frac{1}{K}$ with $b < 1$, which ensures that the right-hand side is U-shaped. Since the LHS is upward sloping and the RHS is a U-shaped parabola centered at 0, if the roots are real there is always a positive one. In order to ensure there are real roots, we need:

$$\begin{aligned}
(K(2\bar{x}^2(1 - \varepsilon)\varepsilon))^2 + 4\bar{x}^4(1 - \varepsilon)(K(1 - \varepsilon) - 1)\varepsilon(1 - K\varepsilon) &\geq 0 \\
K^2 4\bar{x}^4(1 - \varepsilon)^2\varepsilon^2 + 4\bar{x}^4(1 - \varepsilon)(K(1 - \varepsilon) - 1)\varepsilon(1 - K\varepsilon) &\geq 0 \\
K^2(1 - \varepsilon)\varepsilon + (K(1 - \varepsilon) - 1)(1 - K\varepsilon) &\geq 0 \\
K^2(1 - \varepsilon)\varepsilon + (K(1 - \varepsilon) - 1 - \varepsilon K^2(1 - \varepsilon) + K\varepsilon) &\geq 0 \\
K^2(1 - \varepsilon)\varepsilon + (K - 1 - K^2\varepsilon(1 - \varepsilon)) &\geq 0 \\
K - 1 &\geq 0,
\end{aligned}$$

which always holds since the kurtosis is always greater than one.

OA.C Additional Proofs

OA.C.1 Propositions OA.1 and OA.2

The value function in the inaction region is given by $V(z) = -\frac{\sigma^2}{\rho^2}B - \frac{B}{\rho}z^2 + 2b \cosh(\lambda z)$, where $\lambda = \sqrt{2\rho/\sigma^2}$. We need to find the optimal inaction threshold, call it $\bar{x}$. By symmetry, we only need to focus on one side, and we know that the maximum of $V(z)$ is attained at 0. Assume the system we want to solve is given by:

$$\begin{aligned}
0 &= -B\frac{x^2}{\rho} + 2b(\cosh(\lambda x) - 1) + c(\omega, \theta, b, x) \\
0 &= -2B\frac{x}{\rho} + 2b\lambda \sinh(\lambda x),
\end{aligned}$$

where the adjustment cost has the form:

$$c(\omega, \theta, b, x) = -\theta(1 - \omega) \log \left(\int_{-\infty}^{\infty} \bar{f}(z) \exp((V(z) - V(0)) / \theta((1 - \omega))) dz \right) - \frac{\theta \log(\theta)}{4} \omega,$$

with $\omega \in (0, 1]$. Here, in order to adjust, you must pay a fixed cost $-\frac{\theta \log(\theta)}{4}\omega$, and then you have to choose a distribution subject to a control cost $\theta(1 - \omega) \int \log(f(x)/\bar{f}(x))f(x)dx$. We are interested in approximating the solution when $\theta \rightarrow 0$. Using the second equation to solve out b , we can write this as a single system with a single unknown:

$$0 = -B\frac{x^2}{\rho} + \frac{2Bx}{\rho\lambda \sinh(\lambda x)}(\cosh(\lambda x) - 1) + c(\omega, \theta, \frac{Bx}{\rho\lambda \sinh(\lambda x)}, x). \quad (\text{OA.C.1})$$

APPROXIMATION. Guess the solution has the form:

$$x = (A \times \theta |\log \theta|)^{1/4} + o((\theta |\log \theta|)^{1/2}),$$

with an A to be determined. Approximating $b(x) = \frac{2Bx}{\rho\lambda \sinh(\lambda x)}$, since $\sinh(\lambda x) = \lambda x + \frac{\lambda^3 x^3}{6} + O(x^5)$ as $x \rightarrow 0$,

$$b(x) = \frac{Bx}{\rho\lambda(\lambda x + \frac{\lambda^3 x^3}{6} + O(x^5))} = \frac{B}{\rho\lambda^2} \cdot \frac{1}{1 + \frac{\lambda^2 x^2}{6} + O(x^4)}.$$

Expanding the reciprocal gives

$$b(x) = \frac{B}{\rho\lambda^2} \left(1 - \frac{\lambda^2 x^2}{6} + O(x^4)\right) = \frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho} + O(x^4),$$

Then from the expression of the coefficients of the value function:

$$b = \frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho} + o((\theta |\log \theta|)^{1/2}).$$

Thus the value function in the continuation region is:

$$\begin{aligned} V(z) &= -B\frac{\sigma^2}{\rho^2} - B\frac{z^2}{\rho} + 2\left(\frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho}\right)(1 + \lambda^2\frac{z^2}{2} + \lambda^4\frac{z^4}{24}) + o(z^4) \\ &= V(0) - B\frac{z^2}{\rho} + 2\left(\frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho}\right)\lambda^2\frac{z^2}{2} + 2\left(\frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho}\right)\lambda^4\frac{z^4}{24} + o(z^4) \\ &= V(0) - Bz^2\left[\frac{1}{\rho} - \frac{1}{\rho} + \frac{x^2}{6\rho}\lambda^2\right] + 2\left(\frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho}\right)\lambda^4\frac{z^4}{24} + o(z^4) \\ &= V(0) - Bz^2\left[\frac{x^2}{6\rho}\lambda^2\right] + \left(\frac{B}{\rho\lambda^2} - B\frac{x^2}{6\rho}\right)\lambda^4\frac{z^4}{12} + o(z^4), \end{aligned}$$

thus

$$\begin{aligned}
\frac{V(z) - V(0)}{\theta(1 - \omega)} &= -\frac{Bz^2}{\theta(1 - \omega)} \left[\frac{\bar{x}^2}{6\rho} \lambda^2 \right] + 2 \left(\frac{1}{2\rho\lambda^2} - \frac{\bar{x}^2}{12\rho} \right) \lambda^4 \frac{Bz^4}{12\theta(1 - \omega)} + o(z^4/\theta) \\
&= -\frac{Bz^2}{\theta(1 - \omega)} \left[\frac{\sqrt{A}}{6\rho} \lambda^2 \right] \sqrt{\theta|\log \theta|} + 2 \left(\frac{1}{2\rho\lambda^2} - \frac{\sqrt{A}\sqrt{\theta|\log \theta|}}{12\rho} \right) \lambda^4 \frac{Bz^4}{12\theta(1 - \omega)} + o(z^4/\theta) \\
&= -\frac{1}{2(1 - \omega)} \left(\frac{z}{(\theta/|\log(\theta)|)^{1/4}} \right)^2 \left[\frac{2B\sqrt{A}}{6\rho} \lambda^2 \right] + B\lambda^2 \frac{(z/\theta^{1/4})^4}{12(1 - \omega)\rho} \\
&\quad - \lambda^4 B \frac{(z/\theta^{1/4})^4}{12(1 - \omega)} \frac{\sqrt{A}\sqrt{\theta|\log \theta|}}{6\rho} + o(z^4/\theta).
\end{aligned}$$

Let $t = \frac{z}{(\theta/|\log(\theta)|)^{1/4}}$. Then $dz = (\theta/|\log(\theta)|)^{1/4} dt$. Thus, changing variables in the relevant integral:

$$\begin{aligned}
\int_{-\infty}^{\infty} \bar{f}(z) \exp((V(z) - V(0))/\theta((1 - \omega))) dz &= (\theta/|\log(\theta)|)^{1/4} \int_{-\infty}^{\infty} \bar{f}((\theta/|\log(\theta)|)^{1/4} t) \times \\
&\quad \exp\left(\left(V((\theta/|\log(\theta)|)^{1/4} t) - V(0)\right)/\theta((1 - \omega))\right) dt \\
&= (\theta/|\log(\theta)|)^{1/4} \int_{-\infty}^{\infty} \bar{f}((\theta/|\log(\theta)|)^{1/4} t) \times \\
&\quad \exp\left(-\frac{t^2}{2} \left[\frac{2B\sqrt{A}}{6\rho(1 - \omega)} \lambda^2 \right] + \lambda^2 \frac{2B(t)^4}{24\rho(1 - \omega)} / |\log(\theta)| \right. \\
&\quad \left. - \lambda^4 \frac{(t)^4}{24} \frac{2B\sqrt{A}}{6\rho(1 - \omega)} \sqrt{\frac{\theta}{|\log(\theta)|}} + o(z^4/\theta) \right) dt.
\end{aligned}$$

As $\theta \rightarrow 0$, $\frac{1}{|\log(\theta)|} \rightarrow 0$ and $\sqrt{\frac{\theta}{|\log(\theta)|}} \rightarrow 0$, thus the only relevant part for the integral is the first quadratic term. Thus:

$$\begin{aligned}
\int_{-\infty}^{\infty} \bar{f}(z) \exp((V(z) - V(0))/\theta((1 - \omega))) dz &= (\theta/|\log(\theta)|)^{1/4} \bar{f}(0) \times \\
&\quad \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2} \left[\frac{2B\sqrt{A}}{6\rho(1 - \omega)} \lambda^2 \right] \right) dz + o((\theta/|\log(\theta)|)^{1/4}),
\end{aligned}$$

where the integral can be evaluated because it is Gaussian. Let

$$K = \bar{f}(0) \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2} \left[\frac{2B\sqrt{A}}{6\rho(1 - \omega)} \lambda^2 \right] \right) dt.$$

Substituting into the expression for c ,

$$c(\omega, \theta, b, x) = -(1 - \omega)\theta \log(\theta)/4 - \theta(1 - \omega) \log(1/|\log(\theta)|)/4 + o(\theta|\log(\theta)|) - \frac{\theta \log(\theta)}{4} \omega,$$

since the terms with K are scaled by θ and thus go to zero faster than $\theta|\log(\theta)|$. However, note that:

$$\begin{aligned}\lim_{\theta \rightarrow 0} \frac{\theta \log(1/|\log(\theta)|)}{\theta \log(\theta)} &= -\lim_{\theta \rightarrow 0} \frac{\log(|\log(\theta)|)}{\log(\theta)} \\ &=^{L'H} -\lim_{\theta \rightarrow 0} \frac{\frac{1}{|\log(\theta)|} \frac{1}{\theta}}{1/\theta} \\ &= -\lim_{\theta \rightarrow 0} \frac{1}{|\log(\theta)|} = 0,\end{aligned}$$

and thus we can lump that extra term in the $o(\theta|\log(\theta)|)$ part. Therefore, the correct approximation of the cost is:

$$c(\omega, \theta, b, x) = \frac{\theta|\log(\theta)|}{4} + o(\theta|\log(\theta)|).$$

On the other hand, expanding the side of the equation with x to fourth order gives:

$$\begin{aligned}-B\frac{x^2}{\rho} + \frac{2Bx}{\rho\lambda \sinh(\lambda x)}(\cosh(\lambda x) - 1) &= -B\frac{x^2}{\rho} + \frac{2B}{\rho\lambda}x \frac{\lambda^2 \frac{x^2}{2} + \lambda^4 \frac{x^4}{24} + O(x^6)}{\lambda x + \lambda^3 \frac{x^3}{6} + O(x^5)} \\ &= -B\frac{x^2}{\rho} + \frac{2B}{\rho\lambda^2} \frac{\lambda^2 \frac{x^2}{2} + \lambda^4 \frac{x^4}{24} + O(x^6)}{1 + \lambda^2 \frac{x^2}{6} + O(x^4)}.\end{aligned}$$

Using that for z close to zero $\frac{1}{1+z} = 1 - z + O(z^2)$ we have:

$$\begin{aligned}-B\frac{x^2}{\rho} + \frac{2Bx}{\rho\lambda \sinh(\lambda x)}(\cosh(\lambda x) - 1) &= -B\frac{x^2}{\rho} + \frac{B}{\rho} \left(x^2 + \lambda^2 \frac{x^4}{12} + O(x^6) \right) \left(1 - \lambda^2 \frac{x^2}{6} + O(x^4) \right) \\ &= -B\frac{x^2}{\rho} + \frac{B}{\rho} \left(x^2 + \lambda^2 \frac{x^4}{12} - \lambda^2 \frac{x^4}{6} + O(x^6) \right) \\ &= -\frac{B}{\rho} \lambda^2 \frac{x^4}{12} + O(x^6),\end{aligned}$$

thus, equating the leading terms:

$$\begin{aligned}\frac{\theta|\log(\theta)|}{4} &= \frac{B}{\rho} \lambda^2 \frac{x^4}{12} \\ x &= \left(\frac{12\rho}{B\lambda^2} \times \left(\frac{\theta|\log(\theta)|}{4} \right) \right)^{1/4}\end{aligned}$$

$$x = \left(\frac{6\sigma^2}{B} \times \left(\frac{\theta |\log(\theta)|}{4} \right) \right)^{1/4},$$

which confirms the guess with $A = \frac{6\sigma^2}{4B}$, which is exactly the formula in Dixit (1991) for a (small) menu cost $c = \frac{\theta |\log(\theta)|}{4}$. Thus, under this parametrization, the composition does not matter: we always generate an inaction region as if the cost were $\frac{\theta |\log(\theta)|}{4}$. The error bound follows from an argument similar to Dixit (1991).

ASYMPTOTIC DISTRIBUTION. Given this, we can derive the asymptotic distribution. Given our approximation of the value function, we know the asymptotic distribution will have to scale with $(\theta/|\log(\theta)|)^{1/4}$. Looking at the resulting expression we have that:

$$\begin{aligned} f(x) &\propto \exp\left(-\frac{t^2}{2} \left[\frac{2B\sqrt{A}}{6\rho(1-\omega)} \lambda^2 \right] \right) \\ &= \exp\left(-\frac{t^2}{2} \frac{1}{(1-\omega)\sqrt{\frac{3\sigma^2}{2B}}} \right), \end{aligned}$$

thus, formally, if x^* denotes the reset gap, we have

$$(\theta/|\log(\theta)|)^{-1/4} x^* \rightarrow^d \mathcal{N}(0, (1-\omega)\sqrt{\frac{3\sigma^2}{2B}}),$$

so from here we see that the asymptotic variance is increasing in $1-\omega$ (the share of total cost that is control costs) and σ^2 , and it is decreasing in B (if there is higher curvature of the profit function, mistakes are more costly, and thus the firm makes smaller mistakes in price space). A more friendly way to write this is that, along the limit $x = \left(\frac{6\sigma^2}{B} \times \left(\frac{\theta |\log(\theta)|}{4} \right) \right)^{1/4}$, so

$$x^* \approx^d \mathcal{N}\left(0, \frac{3(1-\omega)\sigma^2\theta}{2B\bar{x}^2}\right),$$

from here we can see that the composition of the costs does have an effect on the asymptotic distribution: a higher share of menu cost (ω close to 1) makes the asymptotic distribution less dispersed. Yet a third way of writing it is in $x/\bar{x}$ terms. In particular, the above way of writing it implies that:

$$|\log(\theta)|^{1/2} \frac{x^*}{\bar{x}} \rightarrow^d \mathcal{N}(0, (1-\omega))$$

$$\frac{x^*}{\bar{x}} \approx^d \mathcal{N}(0, \frac{(1-\omega)}{|\log(\theta)|}),$$

where we used that $x^4 = \frac{6\sigma^2}{4B}|\theta \log(\theta)|$ and simplified. So we do get a “square root asymptotics”, but it is in $\frac{x^*}{\bar{x}}$ space, and since $\bar{x}$ also moves, the correct overall scaling with θ is the one mentioned above. This also gives some intuition for the result below that the uncertainty about the reset point does not matter at order $\theta|\log(\theta)|$: since the asymptotic distribution converges to a point mass at zero faster than the bands, it cannot be the case that the uncertainty about the reset point matters (at this order) for the determination of the bands, since the bands converge more slowly than the reset price distribution. However, the convergence rate is $|\log(\theta)|^{1/2}$, which is quite slow, so for small but not extremely small θ the asymptotic distribution will be far from degenerate.

EXPECTED VALUE OF RESETTING. Given that we have the asymptotic distribution, we can also try to compute the expected value upon resetting:

$$\mathbb{E}[V] = \int V(x)f(x)dx.$$

Using the approximation to the value function, $V(x) = V^{\max} + (1-\omega)\theta \frac{V(x) - V^{\max}}{(1-\omega)\theta}$ we have:

$$\begin{aligned} \mathbb{E}[V] &= V^{\max} - (1-\omega)\theta \int \frac{1}{2} \left(\frac{z}{(\theta/|\log(\theta)|)^{1/4}} \right)^2 \left[\frac{\sqrt{A}}{6\rho(1-\omega)} \lambda^2 \right] \frac{\sqrt{\frac{\sqrt{A}}{6\rho(1-\omega)} \lambda^2}}{\sqrt{2\pi}(\theta/|\log(\theta)|)^{1/2}} \times \\ &\quad \exp\left(-\frac{1}{2} \left(\frac{z}{(\theta/|\log(\theta)|)^{1/4}} \right)^2 \left[\frac{\sqrt{A}}{6\rho(1-\omega)} \lambda^2 \right] \right) dz. \end{aligned}$$

Letting $y = \frac{z}{(\theta/|\log(\theta)|)^{1/4}} \times \left[\frac{\sqrt{A}}{6\rho(1-\omega)} \lambda^2 \right]^{1/2}$, then $dz = (\theta/|\log(\theta)|)^{1/4} \left[\frac{\sqrt{A}}{6\rho(1-\omega)} \lambda^2 \right]^{-1/2} dy$ and thus:

$$\begin{aligned} \mathbb{E}[V] &= V^{\max} - \frac{(1-\omega)\theta}{2} \int y^2 \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \\ &= V^{\max} + o(\theta|\log(\theta)|), \end{aligned}$$

therefore, inaction is a consequence of the cost component of the control cost model: the fact that x is uncertain does not matter at the leading order. Intuitively, given that the order of the errors is small, even though we make errors in individual realizations, they go

to zero fast compared to the curvature of V , and thus resetting with error has no effect on the expected value of V upon reset (to the leading order $\theta|\log(\theta)|$).

OA.C.2 Proof of Proposition 12

Let

$$a_j := \left(1 - \int_{-1}^1 \cos(j\pi(u+1)) d\tilde{F}(u)\right) \left(\frac{2}{j\pi}\right)^2, \quad \lambda := -\frac{N(1 - \mathbb{E}[\tilde{x}^2])\pi^2}{2}.$$

Then

$$Y(t) = \frac{1}{1 - \mathbb{E}[\tilde{x}^2]} \sum_{j=1}^{\infty} a_j e^{\lambda j^2 t}.$$

We first note that $Y(0) = 1$. Indeed, for each $u \in [-1, 1]$, $1 - u^2 = \sum_{j=1}^{\infty} (1 - \cos(j\pi(u+1))) \left(\frac{2}{j\pi}\right)^2$, which is the standard Fourier series of $1 - u^2$ on $[-1, 1]$. Integrating both sides with respect to $d\tilde{F}(u)$ gives $1 - \mathbb{E}[\tilde{x}^2] = \sum_{j=1}^{\infty} a_j$, and therefore $Y(0) = 1$. Hence

$$1 - Y(t) = \frac{1}{1 - \mathbb{E}[\tilde{x}^2]} \sum_{j=1}^{\infty} a_j (1 - e^{\lambda j^2 t}).$$

Substituting the definition of a_j ,

$$\begin{aligned} 1 - Y(t) &= \frac{4}{\pi^2(1 - \mathbb{E}[\tilde{x}^2])} \sum_{j=1}^{\infty} \frac{1 - e^{\lambda j^2 t}}{j^2} \\ &\quad - \frac{4}{\pi^2(1 - \mathbb{E}[\tilde{x}^2])} \int_{-1}^1 \sum_{j=1}^{\infty} \cos(j\pi(u+1)) \frac{1 - e^{\lambda j^2 t}}{j^2} d\tilde{F}(u). \end{aligned}$$

We treat the two terms separately. For the first term, the standard small- t expansion is

$$\sum_{j=1}^{\infty} \frac{1 - e^{\lambda j^2 t}}{j^2} = \sqrt{-\pi\lambda t} + o(\sqrt{t}).$$

Therefore

$$\frac{4}{\pi^2(1 - \mathbb{E}[\tilde{x}^2])} \sum_{j=1}^{\infty} \frac{1 - e^{\lambda j^2 t}}{j^2} = \frac{4\sqrt{-\pi\lambda t}}{\pi^2(1 - \mathbb{E}[\tilde{x}^2])} + o(\sqrt{t}).$$

Since $\lambda = -\frac{N(1 - \mathbb{E}[\tilde{x}^2])\pi^2}{2}$, this becomes

$$2\sqrt{\frac{2Nt}{\pi(1 - \mathbb{E}[\tilde{x}^2])}} + o(\sqrt{t}).$$

It remains to show that the integral term is $o(\sqrt{t})$. Define

$$H_t(u) := \sum_{j=1}^{\infty} \cos(j\pi(u+1)) \frac{1 - e^{\lambda j^2 t}}{j^2}.$$

We need to show that $\int_{-1}^1 H_t(u) d\tilde{F}(u) = o(\sqrt{t})$. First, fix any $\delta > 0$. If $|u| \leq 1 - \delta$, then define the partial sums

$$A_m(u) = \sum_{j=1}^m \cos(j\pi(u+1)).$$

These are bounded uniformly in m and $|u| \leq 1 - \delta$. In particular, let $\theta = \pi(u+1)$. Then we have $A_m(u) = \sum_{j=1}^m \cos(j\theta)$. Note that using the formula for a geometric sum:

$$\begin{aligned} \sum_{j=1}^m (\cos(j\theta) + i \sin(j\theta)) &= \sum_{j=1}^m e^{ij\theta} \\ &= e^{i\theta} \frac{1 - e^{i\theta m}}{1 - e^{i\theta}} \\ &= e^{i\theta(m+1)/2} \frac{\sin(m\theta/2)}{\sin(\theta/2)} \\ &= (\cos(\theta(m+1)/2) + i \sin(\theta(m+1)/2)) \frac{\sin(m\theta/2)}{\sin(\theta/2)}, \end{aligned}$$

where I used that $1 - e^{ix} = e^{ix/2}(e^{-ix/2} - e^{ix/2}) = -2ie^{ix/2} \sin(x/2)$. Considering the real part only we have:

$$\sum_{j=1}^m \cos(j\theta) = \cos(\theta(m+1)/2) \frac{\sin(m\theta/2)}{\sin(\theta/2)} \leq \frac{1}{|\sin(\theta/2)|},$$

but since $\sin(\theta/2) = \cos(\frac{\pi u}{2})$ this implies that $|A_m(u)| \leq \frac{1}{\cos(\pi u/2)}$ so $\sup_{m, |u| \leq 1-\delta} |A_m(u)| \leq \sup_{|u| \leq 1-\delta} \frac{1}{\cos(\pi u/2)} = \frac{1}{\cos(\pi(1-\delta)/2)} = C_\delta$. Now let $b_j(t) = e^{\lambda j^2 t}$. Summing by parts:

$$\sum_{j=1}^{\infty} b_j(t) \cos(j\pi(u+1)) = \sum_{j=1}^{\infty} A_j(u)(b_j(t) - b_{j+1}(t)).$$

Then

$$\left| \sum_{j=1}^{\infty} b_j(t) \cos(j\pi(u+1)) \right| \leq \sup |A_j(u)| \sum_{j=1}^{\infty} (b_j(t) - b_{j+1}(t)) \leq \sup |A_j(u)| b_1(t) \leq C_\delta.$$

Now note that for each finite m

$$\sum_{j=1}^m \cos(j\pi(u+1)) \frac{1 - e^{\lambda j^2 t}}{j^2} = -\lambda \int_0^t \sum_{j=1}^m \cos(j\pi(u+1)) e^{\lambda j^2 s} ds,$$

since $-\lambda \int_0^t e^{\lambda j^2 s} ds = \frac{1 - e^{\lambda j^2 t}}{j^2}$. Using the uniform bound found above:

$$\left| \sum_{j=1}^m \cos(j\pi(u+1)) \frac{1 - e^{\lambda j^2 t}}{j^2} \right| \leq -\lambda \int_0^t C_\delta ds = C'_\delta t,$$

so letting $m \rightarrow +\infty$ we obtain that $|H_t(u)| \leq C'_\delta t$ so $|H_t(u)| = O(t)$ for all $|u| \leq 1 - \delta$ as claimed. Now use the assumption that $\tilde{F}$ has no atoms at -1 or 1 . Given $\varepsilon > 0$, choose $\delta > 0$ so small that $\tilde{F}(\{|u| > 1 - \delta\}) < \varepsilon$. Then split the integral:

$$\int_{-1}^1 H_t(u) d\tilde{F}(u) = \int_{|u| \leq 1 - \delta} H_t(u) d\tilde{F}(u) + \int_{|u| > 1 - \delta} H_t(u) d\tilde{F}(u).$$

For the first part, we've shown that

$$\left| \int_{|u| \leq 1 - \delta} H_t(u) d\tilde{F}(u) \right| \leq C'_\delta t.$$

For the second part,

$$\left| \int_{|u| > 1 - \delta} H_t(u) d\tilde{F}(u) \right| \leq C\sqrt{t} \tilde{F}(\{|u| > 1 - \delta\}) \leq C\varepsilon\sqrt{t}.$$

Uniformly in $u \in [-1, 1]$, $|H_t(u)| \leq \sum_{j=1}^\infty \frac{1 - e^{\lambda j^2 t}}{j^2} = O(\sqrt{t})$.

$$\left| \int_{-1}^1 H_t(u) d\tilde{F}(u) \right| \leq C_\delta t + C\varepsilon\sqrt{t}.$$

Divide by $\sqrt{t}$ and let $t \downarrow 0$. This yields

$$\limsup_{t \downarrow 0} \frac{1}{\sqrt{t}} \left| \int_{-1}^1 H_t(u) d\tilde{F}(u) \right| \leq C\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\int_{-1}^1 H_t(u) d\tilde{F}(u) = o(\sqrt{t}).$$

Combining the two parts, we obtain:

$$Y(t) = 1 - 2\sqrt{\frac{2Nt}{\pi(1 - \mathbb{E}[\tilde{x}^2])}} + o(\sqrt{t}).$$

as claimed.

OA.C.3 Invariance to Small Inflation Trend

The proof follows the same symmetry arguments as Alvarez et al. (2016a), so I keep it relatively brief. The problem with trend inflation requires additional state variables. For the manager, we can have the price gaps of individual products x_i as states, which now follow $dx_i = -\mu dt + \sigma dW_i$ independently. The HJBQVI of this problem is:

$$0 = \min \left\{ \rho V - \left(\sum_i (-Bx_i^2 - \mu V_i + \frac{\sigma^2}{2} V_{ii}) + \kappa \int_0^\Psi \max\{V^*(0) - (\psi + V), 0\} dG(\psi) \right), V - (V^*(\Psi)) \right\},$$

where $V_i = \frac{\partial V}{\partial x_i}$. The max inside the integral operator indicates that the firm only resets if doing so increases value. The value of resetting $V^*(0)$ is defined as:

$$V^*(0) = \max_{\{\tilde{f}_i\}} \int \dots \int V(x_1, \dots, x_n) \prod_{i=1}^n \tilde{f}_i(x_i) dx_i - \theta \sum_{i=1}^n \int \log(\tilde{f}_i(x_i)/\bar{f}(x_i)) \tilde{f}_i(x_i) dx_i \quad \text{s.t.} \int \tilde{f}_i(x_i) dx_i = 1 \quad \forall i.$$

First, note that if $\bar{f}$ is symmetric around 0, the cost of choosing $\hat{f}(x) = f(-x)$ is the same as choosing f , since:

$$\begin{aligned} \int_{-\infty}^{\infty} \log(f(-x)/\bar{f}(x)) f(-x) dx &= \int_{-\infty}^{\infty} \log(f(u)/\bar{f}(-u)) f(u) du \\ &= \int_{-\infty}^{\infty} \log(f(u)/\bar{f}(u)) f(u) du, \end{aligned}$$

where we just change variables and use symmetry of $\bar{f}$. Given this, we start by proving some symmetry properties of $V(x_1, \dots, x_n; \mu)$ and $f(x; \mu)$.

Lemma OA.1. $V(x_1, \dots, x_n; -\mu) = V(-x_1, \dots, -x_n; \mu)$, $V^*(0; \mu) = V^*(0; -\mu)$ and $f(x, \mu) = f(-x, -\mu)$.

Proof. Define $U(u_1, \dots, u_n; \mu) := V(-u_1, \dots, -u_n; \mu)$, so that $U_i = -V_i(-u; \mu)$ and $U_{ii} = V_{ii}(-u; \mu)$ for all i . Substituting $x_i = -u_i$ into the original HJBQVI and using these relations

yields:

$$0 = \min \left\{ \rho U - \left(\sum_i \left(-Bu_i^2 + \mu U_i + \frac{\sigma^2}{2} U_{ii} \right) + \kappa \int_0^{\bar{\Psi}} \max \{U^*(0) - (\psi + U), 0\} dG(\psi) \right), U - (U^*(\Psi)) \right\},$$

where the sign on μU_i has flipped relative to the original equation. This is precisely the HJBQVI for an economy with drift parameter $-\mu$. It remains to verify that the reset value is self-consistent, i.e., $U^*(0; \mu) = V^*(0; \mu)$. By definition:

$$U^*(0; \mu) = \max_{\{\tilde{f}_i\}} \int \cdots \int V(-u_1, \dots, -u_n; \mu) \prod_i \tilde{f}_i(u_i) du - \theta \sum_i \int \log \left(\frac{\tilde{f}_i(u_i)}{\tilde{f}(u_i)} \right) \tilde{f}_i(u_i) du_i.$$

Change variables $y_i = -u_i$ and define $\hat{f}_i(y_i) = \tilde{f}_i(-y_i)$, so that $\int \hat{f}_i(y_i) dy_i = 1$. The entropy cost is thus:

$$\int \log \left(\frac{\tilde{f}_i(-y_i)}{\tilde{f}(-y_i)} \right) \tilde{f}_i(-y_i) dy_i = \int \log \left(\frac{\hat{f}_i(y_i)}{\hat{f}(-y_i)} \right) \hat{f}_i(y_i) dy_i = \int \log \left(\frac{\hat{f}_i(y_i)}{\hat{f}(y_i)} \right) \hat{f}_i(y_i) dy_i,$$

where the last equality uses the symmetry of $\bar{f}$ around zero. Therefore:

$$U^*(0; \mu) = \max_{\{\hat{f}_i\}} \int \cdots \int V(y; \mu) \prod_i \hat{f}_i(y_i) dy - \theta \sum_i \int \log \left(\frac{\hat{f}_i(y_i)}{\hat{f}(y_i)} \right) \hat{f}_i(y_i) dy_i = V^*(0; \mu).$$

Thus $U(u; \mu)$ satisfies the HJBQVI for the $(-\mu)$ -economy with reset value $V^*(0; \mu)$. By uniqueness of the solution to the HJBQVI, $U(u; \mu) = V(u; -\mu)$, which gives $V(-x; \mu) = V(x; -\mu)$, or equivalently $V(x_1, \dots, x_n; -\mu) = V(-x_1, \dots, -x_n; \mu)$. In order to see that $V^*(0; -\mu) = V^*(0; \mu)$, take the definition of $V^*(0; \mu)$ and apply the first identity to write $V(x; \mu) = V(-x; -\mu)$, the same change of variables $y_i = -x_i$ and $\hat{f}_i(y_i) = \tilde{f}_i(-y_i)$ (again invoking symmetry of $\bar{f}$) yields:

$$V^*(0; \mu) = \max_{\{\hat{f}_i\}} \int \cdots \int V(y; -\mu) \prod_i \hat{f}_i(y_i) dy - \theta \sum_i \int \log \left(\frac{\hat{f}_i(y_i)}{\hat{f}(y_i)} \right) \hat{f}_i(y_i) dy_i = V^*(0; -\mu).$$

Finally, the change of variables above defines a bijection between the feasible sets of the optimization problems for μ and $-\mu$:

$$\tilde{f}_i(\cdot; \mu) \longleftrightarrow \hat{f}_i(\cdot; -\mu) = \tilde{f}_i(-\cdot; \mu),$$

that preserves the objective value exactly. Since the KL divergence term $\theta \sum_i \int \log(\tilde{f}_i/\bar{f}) \tilde{f}_i dx_i$ is strictly convex in $\tilde{f}_i$, the optimizer is unique. Therefore, if $f^*(\cdot; \mu)$ is the optimal posterior for the case with μ , then $\hat{f}_i^*(y) := f_i^*(-y; \mu)$ must be the unique optimizer for the economy with $-\mu$, giving $f(x; \mu) = f(-x; -\mu)$. $\square$

The fact that $f(x, \mu) = f(-x, -\mu)$ implies that all even centered moments of x^* satisfy $\mathbb{E}[(x^* - \mathbb{E}[x^*])^m | \mu] = \mathbb{E}[(x^* - \mathbb{E}[x^*])^m | -\mu]$. Once we have that the value function is symmetric, the generalized hazard function $\Lambda(x_1, \dots, x_n, \mu) = \kappa G(V^*(0; \mu) - V(x; \mu))$ also satisfies $\Lambda(x_1, \dots, x_n, \mu) = \Lambda(-x_1, \dots, -x_n, -\mu)$, and the same holds for the inaction set if $\Psi < +\infty$. In conjunction with $f(x, \mu) = f(-x, -\mu)$, this implies that the probability law for $x_{i,t}$ with μ has exactly the same stochastic process as $u_{i,t} = -x_{i,t}$ with $-\mu$. Given that price changes are $\Delta p = x^* - x$, and since the distribution of x^* under μ is the same as $-x^*$ under $-\mu$, and the same applies to x , we have that the distribution of Δp under μ is the same as the distribution of $-\Delta p$ under $-\mu$. For the same reason as for x^* , this implies that all centered even moments are symmetric around $\mu = 0$, i.e., $\text{Var}(\Delta p | \mu) = \text{Var}(\Delta p | -\mu)$, $\text{Kurt}(\Delta p | \mu) = \text{Kurt}(\Delta p | -\mu)$, and so on. Finally, since the law of $u_{i,t}$ is the same as that of $x_{i,t}$, the invariant density satisfies $g(x_1, \dots, x_n; \mu) = g(-x_1, \dots, -x_n; -\mu)$. Thus, the frequency of price changes N is also symmetric in μ , $N(\mu) = N(-\mu)$. Since all relevant moments for the sufficient statistics result are symmetric around μ , and for any symmetric function ϕ we have that $\phi'(0) = 0$, this shows that all moments are invariant to μ at $\mu = 0$ if we assume that these moments are differentiable as a function of μ . Since the reset gap distribution satisfies those symmetry properties, it follows that β_1 and β_2 are also invariant to first order.

For the output IRF, once we have that $g(x_1, \dots, x_n; \mu) = g(-x_1, \dots, -x_n; -\mu)$ and that the hazard function and inaction sets satisfy the relevant symmetry properties, the same argument in the Appendix of Alvarez et al. (2016a) applies, and thus $\mathcal{M}(\delta, \mu) = -\mathcal{M}(-\delta, -\mu)$, so $\mathcal{M}'(0, \mu) = \mathcal{M}'(0, -\mu)$, and thus once again, non-neutrality is invariant to μ when $\mu = 0$.

OA.C.4 Proof of Calvo Hazard

First, using that the random menu costs have a mass point at zero and that $\Psi \rightarrow +\infty$. For the purpose of this subsection and the next, let $y = \frac{1}{n} \sum_{i=1}^n x_i^2$ be the average squared gap. This alternative convention allows us to see more easily what happens as $n \rightarrow +\infty$. Then HJB in this case reads:

$$\rho V = -By + \sigma^2 V_y + 2 \frac{\sigma^2}{n} y V_{yy} + \kappa \max\{V^* - V(y), 0\},$$

Note that in the case of perfect reset, $V^* = \max_y V(y)$, so the $\max\{V^* - V(y), 0\} = V^* - V(y)$ and we get the typical Calvo HJB. In the case of $n \rightarrow +\infty$, $y^* = \sum_{i=1}^n (x_i)^2/n \rightarrow \text{Var}(x^*)$. Since we are taking $\theta \rightarrow +\infty$, this implies we choose $f^* = \bar{f}$ and thus $V^* = \mathbb{E}[V(y^*)]$.⁷

Using the continuous mapping theorem, we have that as $n \rightarrow +\infty$, $V^* = V(\text{Var}(x^*))$. Now, given the definition of y , its law of motion for an arbitrary n is given by $dy = \sigma^2 dt + 2\sigma\sqrt{y/nd}W$. As $n \rightarrow +\infty$, $dy = \sigma^2 dt$. This implies that upon reset, y starts at $\text{Var}(x^*)$ and deterministically increases. Since $V(y)$ is decreasing in y , this implies that $V(y) \leq V^*$ for all $y \geq \text{Var}(x^*)$. But since y can only be above $\text{Var}(x^*)$ in the continuation region, this implies that $\max\{V^* - V(y), 0\} = V^* - V(y)$ for all relevant y , and thus we get a pure Calvo hazard for the relevant region.

Note that for an arbitrary, finite θ , there will be in general some values for which $V(y) > V^*$. This is because $V^* = \mathbb{E}[V(y^*)] - \theta \mathcal{C}(f^* || \bar{f})$, so the “pure cost” component $\theta \mathcal{C}(f^* || \bar{f})$ will also appear. In that case, the model will behave similarly to a random menu cost model where the cost of resetting if an opportunity arrives is not zero. Note this applies even if $n \rightarrow +\infty$ and the reset point y^* is deterministic: given that the variance of x^* is costly, the firm will have incentives to reduce it and thus will pay a positive cost for it. But this would imply that there is a region where $V(y) > V^*$, thus generating an inaction region.

OA.C.5 Calvo Model with Mistakes and Mean Reversion

OA.C.5.1 Optimality of a Constant Hazard

Consider a firm with profits $\pi = -By$ where as in the previous subsection $y = \frac{1}{n} \sum_{i=1}^n x_i^2$, with $x_i = p_i - z_i$, and $dz_i = -\eta z_i dt + \sigma dW_i$ and so $dx_i = \eta z_i dt - \sigma dW_i$, and the Brownian shocks of each product are independent. As before, whenever prices are reset, I impose that all prices must be reset at the same time. Furthermore, assume as above that $\theta \rightarrow +\infty$ so all reset imperfections are drawn from an exogenous distribution F . In order to get a recursive representation of the problem, note that profits evolve as:

$$\begin{aligned} dy &= \frac{1}{n} \sum_{i=1}^n 2x_i(\eta z_i dt - \sigma dW_i) + \sigma^2 dt \\ &= (2\eta w + \sigma^2)dt - 2\sigma\sqrt{y/nd}W_y, \end{aligned}$$

⁷Alternatively, one can think that the cost functional is $\mathcal{C}(F) = 0$ if $F = \bar{F}$, and $+\infty$ otherwise.

now I have to deal with the $w = \frac{1}{n} \sum_{i=1}^n x_i z_i = \frac{1}{n} \sum_{i=1}^n (p_i - z_i) z_i$ term. Then:

$$\begin{aligned} dw &= \frac{1}{n} \sum_{i=1}^n (p_i - 2z_i)(-\eta z_i dt + \sigma dW_i) - \sigma^2 dt \\ &= (-\eta(w - h) - \sigma^2)dt + \sigma \sqrt{(y + h - 2w)/n} dW_w, \end{aligned}$$

and thus we need to get the law of motion of a new variable, $h = \frac{1}{n} \sum_{i=1}^n z_i^2$, which satisfies:

$$\begin{aligned} dh &= \frac{1}{n} \sum_{i=1}^n 2z_i(-\eta z_i dt + \sigma dW_i) + \sigma^2 dt \\ &= (\sigma^2 - 2\eta h)dt + 2\sigma \sqrt{h/n} dW_h, \end{aligned}$$

completing the other two yields:

$$\begin{aligned} dy &= (2\eta w + \sigma^2)dt - 2\sigma \sqrt{y/n} dW_y \\ dw &= (\eta(h - w) - \sigma^2)dt + \sigma \sqrt{(y + h - 2w)/n} dW_w \\ dh &= (\sigma^2 - 2\eta h)dt + 2\sigma \sqrt{h/n} dW_h. \end{aligned}$$

For finite n , the three Brownian drivers are correlated because they are constructed from the same product shocks. Where the denominators are positive, define

$$dW_y = \frac{\sum_i x_i dW_i}{\sqrt{ny}}, \quad dW_w = \frac{\sum_i (x_i - z_i) dW_i}{\sqrt{n(y + h - 2w)}}, \quad dW_h = \frac{\sum_i z_i dW_i}{\sqrt{nh}}.$$

Equivalently, the diffusion terms in (dy, dw, dh) are

$$\frac{\sigma}{n} \sum_{i=1}^n (-2x_i, x_i - z_i, 2z_i) dW_i,$$

which remains well defined when a denominator above is zero. As in the main text, ε_i^* is the independent, mean-zero reset error, whereas the total post-reset gap is $x_i^* = (\bar{c} - 1)z_i + \varepsilon_i^*$. Taking $n \rightarrow \infty$ we have the deterministic system:

$$\begin{aligned} \dot{y} &= (2\eta w + \sigma^2) \\ \dot{w} &= (\eta(h - w) - \sigma^2) \\ \dot{h} &= (\sigma^2 - 2\eta h). \end{aligned}$$

Assume the product-level desired prices are initially independent draws from their stationary Ornstein–Uhlenbeck distribution, with variance $\sigma^2/(2\eta)$. In the large-product limit, h is constant and equal to the unconditional variance of an OU process, $h = \frac{\sigma^2}{2\eta}$. This reduces the problem to:

$$\begin{aligned}\dot{y} &= (2\eta w + \sigma^2) \\ \dot{w} &= -\eta w - \frac{\sigma^2}{2}.\end{aligned}$$

For any $w(0)$, we can solve the second equation as: $w(t) = -\frac{\sigma^2}{2\eta} + (w(0) + \frac{\sigma^2}{2\eta})e^{-\eta t} = w(0)e^{-\eta t} - \frac{\sigma^2}{2\eta}(1 - e^{-\eta t})$. Given this, note that under our assumptions firms reset to $p_i = \bar{c}z_i + \varepsilon_i^*$ so $w(0) = \frac{1}{n} \sum ((\bar{c} - 1)z_i + \varepsilon_i^*)z_i \rightarrow (\bar{c} - 1)\frac{\sigma^2}{2\eta}$. Since $\bar{c} = \frac{\rho + \kappa}{\rho + \kappa + \eta}$, $\bar{c} - 1 = -\frac{\eta}{\rho + \kappa + \eta}$ and thus $w(t) = -\frac{\sigma^2}{2\eta} \left(\frac{\eta}{\rho + \kappa + \eta} e^{-\eta t} + (1 - e^{-\eta t}) \right)$. Essentially, $w(t)$ starts at $-\frac{\eta}{\rho + \kappa + \eta} \frac{\sigma^2}{2\eta}$ and decreases monotonically and exponentially towards an asymptote of $-\frac{\sigma^2}{2\eta}$. Given that $w > -\frac{\sigma^2}{2\eta}$ this implies that $2\eta w + \sigma^2 > 0$, so $y(t)$ starts at $y(0) = (\bar{c} - 1)^2 \sigma^2 / (2\eta) + \text{Var}(\varepsilon^*)$ and increases monotonically. We can solve for $y(t)$ as:

$$\begin{aligned}y(t) &= y(0) + 2\eta \int_0^t w(s) ds + \sigma^2 t \\ &= y(0) - \sigma^2 \int_0^t \left(-\frac{\rho + \kappa}{\rho + \kappa + \eta} e^{-\eta s} + 1 \right) ds + \sigma^2 t \\ &= y(0) + \sigma^2 \frac{\rho + \kappa}{\rho + \kappa + \eta} \frac{1 - e^{-\eta t}}{\eta}.\end{aligned}$$

In order to see what this implies for profits, I guess and verify that the hazard is indeed constant in the relevant range. In the limiting model considered here, reset is free at an opportunity, so $V^* = V^{**} = V(y(0), w(0))$. From the guess, the value function $V(y, w)$ solves

$$\rho V = -By + (2\eta w + \sigma^2)V_y - \frac{1}{2}(2\eta w + \sigma^2)V_w + \kappa(V^* - V).$$

Given the Calvo hazard, the value function corresponds to the particular solution of the HJB, which is linear in y and w as:

$$V = a_0 y + a_1 w + a_2.$$

Plugging this into the equation we get:

$$\begin{aligned}\rho a_0 &= -B - \kappa a_0 \\ \rho a_1 &= 2\eta a_0 - \eta a_1 - \kappa a_1,\end{aligned}$$

and an equation for a_2 . We can solve for a_0 and a_1 to yield:

$$\begin{aligned}a_0 &= -\frac{B}{\rho + \kappa} \\ a_1 &= -\frac{2\eta B}{(\rho + \kappa)(\rho + \kappa + \eta)}.\end{aligned}$$

This implies that V is monotonically decreasing in both y and w . With that, we can evaluate V at each point in time as:

$$\begin{aligned}V(y(t), w(t)) &= -\frac{B}{\rho + \kappa} \left(y(0) + \sigma^2 \frac{\rho + \kappa}{\rho + \kappa + \eta} \frac{1 - e^{-\eta t}}{\eta} \right) \\ &\quad - \frac{2\eta B}{(\rho + \kappa)(\rho + \kappa + \eta)} \left(-\frac{\sigma^2}{2\eta} \left(\frac{\eta}{\rho + \kappa + \eta} e^{-\eta t} + (1 - e^{-\eta t}) \right) \right) + a_2 \\ &= \sigma^2 B e^{-\eta t} \left(\frac{1}{\eta} \frac{1}{\rho + \kappa + \eta} - \frac{2\eta}{(\rho + \kappa)(\rho + \kappa + \eta)} \frac{1}{2\eta} \frac{\rho + \kappa}{\rho + \kappa + \eta} \right) + \text{Constants}.\end{aligned}$$

but

$$\frac{1}{\eta} \frac{1}{\rho + \kappa + \eta} - \frac{2\eta}{(\rho + \kappa)(\rho + \kappa + \eta)} \frac{1}{2\eta} \frac{\rho + \kappa}{\rho + \kappa + \eta} = \frac{1}{\rho + \kappa + \eta} \left[\frac{1}{\eta} - \frac{1}{\rho + \kappa + \eta} \right] > 0,$$

since $\rho + \kappa > 0$. Thus, upon reset, the value function is monotonically decreasing over time, i.e., $V(y(t), w(t)) < V(y(0), w(0))$ for all $t > 0$, and thus the firm strictly prefers to reset prices if given the chance to do it for free, so a constant hazard is optimal. Therefore, the same justification as in the multiproduct case with $\eta = 0$ works: after the firm resets all its prices, its profits worsen monotonically, and therefore the firm will choose to reset again if given the chance to do it for free.

OA.C.5.2 Aggregation

The above derivation justifies equation (5.2). In this subsection, we derive analytical formulas for the moments of interest. First, I derive the case without mistakes, and later on I add mistakes.

Perfect Reset.

Let $z_n = z_{i,\tau_{i,n}}$ denote the relevant variables at the time of a price change, where we drop the i subscript. The subsequent derivations use the time-series expectations of the relevant variables, but due to ergodicity these are equivalent to the cross-sectional expectations of interest (see Hamilton (1994)). For the unconditional variance of price changes, we have:

$$\begin{aligned}\text{Var}(p_n - p_{n-1}) &= \bar{c}^2 \text{Var}(z_n - z_{n-1}) \\ &= 2\bar{c}^2 (\text{Var}(z_n) - \text{Cov}(z_n, z_{n-1})),\end{aligned}$$

but

$$\begin{aligned}\text{Cov}(z_n, z_{n-1}) &= \mathbb{E}[\mathbb{E}[z_n z_{n-1} | \tau_{i,n}, \tau_{i,n-1}]] \\ &= \mathbb{E}[e^{-\eta \Delta \tau_n} (z_{n-1})^2] \\ &= \text{Var}(z_n) \mathbb{E}[e^{-\eta \Delta \tau_n}],\end{aligned}$$

due to stationarity so

$$\begin{aligned}\text{Var}(p_{i,n} - p_{i,n-1}) &= 2\text{Var}(z) \bar{c}^2 (1 - \frac{\kappa}{\kappa + \eta}) \\ &= \sigma^2 \frac{(\rho + \kappa)^2}{(\rho + \kappa + \eta)^2 (\kappa + \eta)},\end{aligned}$$

when $\eta = 0$, this reduces to σ^2/κ which is the usual formula. Note also that higher η (more mean reversion) decreases the variance. This is intuitive: as the process for flexible prices drifts less far, the size of price changes is smaller.

Throughout this derivation, $\Delta \tau_n = \tau_{i,n} - \tau_{i,n-1}$ denotes a spell duration, and $p_n = p_{i,\tau_{i,n}}$. Durations are independent exponential draws with rate κ , independent of the desired-price shocks and reset errors. Consequently the duration-regression slope is $\tilde{\beta}_3 = 0$, so the adjusted statistic β_2 equals the raw squared-change slope $\tilde{\beta}_2$ in this model. For $\tilde{\beta}_2$ we need to compute:

$$\tilde{\beta}_2 = \frac{\text{Cov}((z_{n+1} - z_n)^2, (z_n - z_{n-1})^2)}{\text{Var}((z_{n+1} - z_n)^2)}.$$

I use the following fact: if (x_1, x_2) are jointly Gaussian with mean-zero, then $\text{Cov}((x_1)^2, (x_2)^2) = 2\text{Cov}(x_1, x_2)^2$. I apply this here conditionally on the stopping times. Since

$$\text{Cov}((z_{n+1} - z_n)^2, (z_n - z_{n-1})^2) = \mathbb{E}[\text{Cov}((z_{n+1} - z_n)^2, (z_n - z_{n-1})^2 | \tau_{n+1}, \tau_n, \tau_{n-1})]$$

$$\begin{aligned}
& + \text{Cov}(\mathbb{E}[(z_{n+1} - z_n)^2 | \tau_{n+1}, \tau_n, \tau_{n-1}], \mathbb{E}[(z_n - z_{n-1})^2 | \tau_{n+1}, \tau_n, \tau_{n-1}]) \\
& = 2\mathbb{E}[\text{Cov}((z_{n+1} - z_n), (z_n - z_{n-1}) | \tau_{n+1}, \tau_n, \tau_{n-1})^2] \\
& = 2\text{Var}(z)^2 \mathbb{E}[(e^{-\eta\Delta\tau_{n+1}} - 1 - e^{-\eta(\Delta\tau_n + \Delta\tau_{n+1})} + e^{-\eta\Delta\tau_n})^2] \\
& = 2\text{Var}(z)^2 \mathbb{E}[(e^{-\eta\Delta\tau_n} - 1)(1 - e^{-\eta\Delta\tau_{n+1}})]^2 \\
& = 2\text{Var}(z)^2 \mathbb{E}[(1 - e^{-\eta\Delta\tau_{n+1}})^2]^2.
\end{aligned}$$

notice that

$$\begin{aligned}
\mathbb{E}[(z_{n+1} - z_n)^2 | \tau_{n+1}, \tau_n, \tau_{n-1}] & = 2(\text{Var}(z) - \text{Cov}(z_{n+1}, z_n)) \\
& = 2\text{Var}(z)(1 - e^{-\eta\Delta\tau_{n+1}}),
\end{aligned}$$

and analogously $\mathbb{E}[(z_n - z_{n-1})^2 | \tau_{n+1}, \tau_n, \tau_{n-1}] = 2\text{Var}(z)(1 - e^{-\eta\Delta\tau_n})$ but since durations are independent, the covariance term is zero. And for the variance:

$$\begin{aligned}
\text{Var}((z_n - z_{n-1})^2) & = 2\mathbb{E}[\text{Cov}((z_n - z_{n-1}), (z_n - z_{n-1}) | \tau_{n+1}, \tau_n, \tau_{n-1})^2] + \text{Var}(\mathbb{E}[(z_n - z_{n-1})^2 | \tau_{n+1}, \tau_n, \tau_{n-1}]) \\
& = 2\text{Var}(z)^2 \mathbb{E}[(2 - 2e^{-\eta\Delta\tau_n})^2] + 4\text{Var}(z)^2 \text{Var}(1 - e^{-\eta\Delta\tau_n}) \\
& = 8\text{Var}(z)^2 \mathbb{E}[(1 - e^{-\eta\Delta\tau_{n+1}})^2] + 4\text{Var}(z)^2 (\mathbb{E}[(1 - e^{-\eta\Delta\tau_n})^2] - \mathbb{E}[1 - e^{-\eta\Delta\tau_n}]^2).
\end{aligned}$$

thus:

$$\begin{aligned}
\tilde{\beta}_2 & = \frac{2\text{Var}(z)^2 (\mathbb{E}[(1 - e^{-\eta\Delta\tau_n})^2])^2}{8\text{Var}(z)^2 \mathbb{E}[(1 - e^{-\eta\Delta\tau_n})^2] + 4\text{Var}(z)^2 \text{Var}(1 - e^{-\eta\Delta\tau_n})} \\
& = \frac{(\mathbb{E}[1 - 2e^{-\eta\Delta\tau_n} + e^{-2\eta\Delta\tau_n}])^2}{4\mathbb{E}[(1 - e^{-\eta\Delta\tau_n})^2] + 2\text{Var}(1 - e^{-\eta\Delta\tau_n})}.
\end{aligned}$$

but $\mathbb{E}[1 - 2e^{-\eta\Delta\tau_{n+1}} + e^{-2\eta\Delta\tau_{n+1}}] = 1 - 2\frac{\kappa}{\kappa + \eta} + \frac{\kappa}{\kappa + 2\eta}$ and $\mathbb{E}[(1 - e^{-\eta\Delta\tau_n})^2] = (1 - \frac{\kappa}{\kappa + \eta})^2 = \frac{\eta^2}{(\kappa + \eta)^2}$.

Note that:

$$\begin{aligned}
\text{Var}((1 - e^{-\eta\Delta\tau_n})) & = 1 - 2\frac{\kappa}{\kappa + \eta} + \frac{\kappa}{\kappa + 2\eta} - \frac{\eta^2}{(\kappa + \eta)^2} \\
& = \frac{(\kappa + \eta)^2(\kappa + 2\eta) - 2\kappa(\kappa + \eta)(\kappa + 2\eta) + \kappa(\kappa + \eta)^2 - \eta^2(\kappa + 2\eta)}{(\kappa + \eta)^2(\kappa + 2\eta)} \\
& = \frac{(\eta - \kappa)(\kappa + \eta)(\kappa + 2\eta) + \kappa(\kappa + \eta)^2 - \eta^2(\kappa + 2\eta)}{(\kappa + \eta)^2(\kappa + 2\eta)} \\
& = \frac{(-\kappa^2)(\kappa + 2\eta) + \kappa(\kappa + \eta)^2}{(\kappa + \eta)^2(\kappa + 2\eta)}
\end{aligned}$$

$$= \frac{\kappa((\kappa + \eta)^2 - \kappa(\kappa + 2\eta))}{(\kappa + \eta)^2(\kappa + 2\eta)} = \frac{\kappa\eta^2}{(\kappa + \eta)^2(\kappa + 2\eta)}.$$

Notice also that this gives us a clean way to compute the kurtosis of price changes, since for any mean zero random variable:

$$\begin{aligned}\text{Var}(x^2) &= \mathbb{E}[x^4] - \mathbb{E}[(x)^2]^2 \\ &= \text{Var}(x)^2(\text{Kurt}(x) - 1).\end{aligned}$$

So:

$$\begin{aligned}\text{Kurt}(\Delta p) &= \text{Kurt}(\Delta z) \\ &= \text{Var}((\Delta z)^2)/(\text{Var}(\Delta z)^2) + 1 \\ &= \frac{8\text{Var}(z)^2(1 - 2\frac{\kappa}{\kappa+\eta} + \frac{\kappa}{\kappa+2\eta}) + 4\text{Var}(z)^2\frac{\kappa\eta^2}{(\kappa + \eta)^2(\kappa + 2\eta)}}{(2\text{Var}(z_n)(\frac{\eta}{\kappa+\eta}))^2} + 1 \\ &= \frac{2(1 - 2\frac{\kappa}{\kappa+\eta} + \frac{\kappa}{\kappa+2\eta}) + \frac{\kappa\eta^2}{(\kappa + \eta)^2(\kappa + 2\eta)}}{((\frac{\eta}{\kappa+\eta}))^2} + 1,\end{aligned}$$

but

$$\begin{aligned}(1 - 2\frac{\kappa}{\kappa + \eta} + \frac{\kappa}{\kappa + 2\eta}) &= \frac{(\kappa + \eta)(\kappa + 2\eta) - 2\kappa(\kappa + 2\eta) + \kappa(\kappa + \eta)}{(\kappa + \eta)(\kappa + 2\eta)} \\ &= \frac{(\eta - \kappa)(\kappa + 2\eta) + \kappa(\kappa + \eta)}{(\kappa + \eta)(\kappa + 2\eta)} \\ &= \frac{\eta^2 - \kappa^2 + \eta(\eta - \kappa) + \kappa(\kappa + \eta)}{(\kappa + \eta)(\kappa + 2\eta)} \\ &= \frac{2\eta^2}{(\kappa + \eta)(\kappa + 2\eta)},\end{aligned}$$

so

$$\begin{aligned}\text{Kurt}(\Delta p) &= \frac{\frac{4\eta^2}{(\kappa + \eta)(\kappa + 2\eta)} + \frac{\kappa\eta^2}{(\kappa + \eta)^2(\kappa + 2\eta)}}{((\frac{\eta}{\kappa+\eta}))^2} + 1 \\ &= \frac{4\eta^2(\kappa + \eta) + \kappa\eta^2}{(\kappa + 2\eta)\eta^2} + 1\end{aligned}$$

$$= \frac{4(\kappa + \eta) + \kappa + (\kappa + 2\eta)}{(\kappa + 2\eta)} = \frac{6(\kappa + \eta)}{\kappa + 2\eta},$$

notice also that:

$$\begin{aligned} \frac{\partial \text{Kurt}(\Delta p)}{\partial \eta} &= \frac{6(\kappa + 2\eta) - 12(\kappa + \eta)}{(\kappa + 2\eta)^2} \\ \frac{\partial \text{Kurt}(\Delta p)}{\partial \eta} &= \frac{-6\kappa}{(\kappa + 2\eta)^2} < 0, \end{aligned}$$

and at $\eta = 0$ this derivative is $-6/\kappa$.

ADDING MISTAKES. First, the variance of price changes:

$$\begin{aligned} \text{Var}(\Delta p_n) &= \bar{c}^2 \text{Var}(z_n - z_{n-1}) + 2\text{Var}(\varepsilon^*) \\ &= \bar{c}^2 \left(\frac{\sigma^2}{\kappa + \eta} \right) + 2\text{Var}(\varepsilon^*), \end{aligned}$$

where we use the value of $\text{Var}(z_n - z_{n-1})$ computed above. For the kurtosis, we need:

$$\begin{aligned} \mathbb{E}[(\Delta p_n)^4] &= \mathbb{E}[(\bar{c}(z_n - z_{n-1}) + \varepsilon_n^* - \varepsilon_{n-1}^*)^4] \\ &= \mathbb{E}[(\bar{c}(z_n - z_{n-1}))^4] + 4\mathbb{E}[(\bar{c}(z_n - z_{n-1}))^3(\varepsilon_n^* - \varepsilon_{n-1}^*)] + \\ &\quad 6\mathbb{E}[(\bar{c}(z_n - z_{n-1}))^2(\varepsilon_n^* - \varepsilon_{n-1}^*)^2] + 4\mathbb{E}[(\bar{c}(z_n - z_{n-1}))(\varepsilon_n^* - \varepsilon_{n-1}^*)^3] \\ &\quad + \mathbb{E}[(\varepsilon_n^* - \varepsilon_{n-1}^*)^4] \\ &= \mathbb{E}[(\bar{c}(z_n - z_{n-1}))^4] + 6\mathbb{E}[(\bar{c}(z_n - z_{n-1}))^2]\mathbb{E}[(\varepsilon_n^* - \varepsilon_{n-1}^*)^2] + \mathbb{E}[(\varepsilon_n^* - \varepsilon_{n-1}^*)^4], \end{aligned}$$

but $\mathbb{E}[(\varepsilon_n^* - \varepsilon_{n-1}^*)^2] = 2\text{Var}(\varepsilon^*)$ and $\mathbb{E}[(\varepsilon_n^* - \varepsilon_{n-1}^*)^4] = 2\mathbb{E}[(\varepsilon^*)^4] + 6\mathbb{E}[(\varepsilon^*)^2]^2 = 2\text{Var}(\varepsilon^*)^2(\text{Kurt}(\varepsilon^*) + 3)$ and we have already computed the other moments above, so we get:

$$\mathbb{E}[(\Delta p_n)^4] = \mathbb{E}[(\bar{c}(z_n - z_{n-1}))^4] + 12\mathbb{E}[(\bar{c}(z_n - z_{n-1}))^2]\text{Var}(\varepsilon^*) + 2\text{Var}(\varepsilon^*)^2(\text{Kurt}(\varepsilon^*) + 3),$$

and with that we can compute the kurtosis. For the predictability coefficients:

$$\begin{aligned} \frac{\text{Cov}(\bar{c}(z_{n+1} - z_n) + \varepsilon_{n+1}^* - \varepsilon_n^*, \bar{c}(z_n - z_{n-1}) + \varepsilon_n^* - \varepsilon_{n-1}^*)}{\text{Var}(\Delta p_n)} &= \\ \frac{\text{Cov}(\bar{c}(z_{n+1} - z_n), \bar{c}(z_n - z_{n-1})) - \text{Var}(\varepsilon^*)}{\text{Var}(\Delta p_n)}. \end{aligned}$$

where we have computed the covariance already. Finally, for $\tilde{\beta}_2$

$$\tilde{\beta}_2 = \frac{\text{Cov}((\bar{c}(z_{n+1} - z_n) + \varepsilon_{n+1}^* - \varepsilon_n^*)^2, (\bar{c}(z_n - z_{n-1}) + \varepsilon_n^* - \varepsilon_{n-1}^*)^2)}{\text{Var}((\Delta p_n)^2)},$$

for the numerator, note that:

$$\begin{aligned} (\bar{c}(z_{n+1} - z_n) + \varepsilon_{n+1}^* - \varepsilon_n^*)^2 &= (\bar{c}(z_{n+1} - z_n))^2 + 2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n) \\ &\quad + (\varepsilon_{n+1}^* - \varepsilon_n^*)^2 \\ &= (\bar{c}(z_{n+1} - z_n))^2 + 2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n) \\ &\quad + (\varepsilon_{n+1}^*)^2 - 2\varepsilon_{n+1}^*\varepsilon_n^* + (\varepsilon_n^*)^2. \end{aligned}$$

so the covariance in the numerator reduces to:

$$\begin{aligned} \text{Cov}(\dots) &= \text{Cov}((\bar{c}(z_{n+1} - z_n))^2, (\bar{c}(z_n - z_{n-1}))^2) + \\ &\quad \text{Cov}((\bar{c}(z_{n+1} - z_n))^2, 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1})) + \\ &\quad \text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), (\bar{c}(z_n - z_{n-1}))^2) + \\ &\quad \text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1})) + \\ &\quad \text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), (\varepsilon_n^* - \varepsilon_{n-1}^*)^2) + \\ &\quad \text{Cov}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2, 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1})) + \\ &\quad \text{Cov}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2, (\varepsilon_n^* - \varepsilon_{n-1}^*)^2). \end{aligned}$$

where I omitted some terms that are clearly zero due to independence. Note that $\text{Cov}((\bar{c}(z_{n+1} - z_n))^2, 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1})) = 0$ since the second part has mean zero, and we can use independence. For the same reason $\text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), (\bar{c}(z_n - z_{n-1}))^2)$ drops out. A similar argument applies to $\text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), (\varepsilon_n^* - \varepsilon_{n-1}^*)^2)$ and $\text{Cov}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2, 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1}))$. So we are left with:

$$\begin{aligned} \text{Cov}(\dots) &= \text{Cov}((\bar{c}(z_{n+1} - z_n))^2, (\bar{c}(z_n - z_{n-1}))^2) + \\ &\quad \text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1})) + \\ &\quad \text{Cov}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2, (\varepsilon_n^* - \varepsilon_{n-1}^*)^2). \end{aligned}$$

We have already computed the first term. For the second term, note that both components have mean zero, so:

$$\begin{aligned}
& \text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), 2(\varepsilon_n^* - \varepsilon_{n-1}^*)\bar{c}(z_n - z_{n-1})) = \\
& 4\bar{c}^2 \text{Cov}((z_{n+1} - z_n), (z_n - z_{n-1})) \text{Cov}(\varepsilon_{n+1}^* - \varepsilon_n^*, \varepsilon_n^* - \varepsilon_{n-1}^*) = \\
& 4\bar{c}^2 \text{Cov}((z_{n+1} - z_n), (z_n - z_{n-1}))(-\text{Var}(\varepsilon^*)),
\end{aligned}$$

and for the last term:

$$\begin{aligned}
\text{Cov}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2, (\varepsilon_n^* - \varepsilon_{n-1}^*)^2) &= \text{Cov}((\varepsilon_{n+1}^*)^2 - 2\varepsilon_{n+1}^*\varepsilon_n^* \\
&\quad + (\varepsilon_n^*)^2, (\varepsilon_n^*)^2 - 2\varepsilon_n^*\varepsilon_{n-1}^* + (\varepsilon_{n-1}^*)^2) \\
&= \text{Var}((\varepsilon_n^*)^2) = \text{Var}(\varepsilon^*)^2[\text{Kurt}(\varepsilon^*) - 1],
\end{aligned}$$

so putting everything together the numerator is:

$$\begin{aligned}
\text{Cov}(\dots) &= \bar{c}^4 \text{Cov}((z_{n+1} - z_n)^2, (z_n - z_{n-1})^2) \\
&\quad - 4\bar{c}^2 \text{Var}(\varepsilon^*) \text{Cov}((z_{n+1} - z_n), (z_n - z_{n-1})) + \text{Var}(\varepsilon^*)^2[\text{Kurt}(\varepsilon^*) - 1].
\end{aligned}$$

For the denominator:

$$\begin{aligned}
\text{Var}((\Delta p)^2) &= \text{Var}((\bar{c}(z_{n+1} - z_n))^2) + \text{Var}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n)) + \text{Var}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2) \\
&\quad + 2\text{Cov}((\bar{c}(z_{n+1} - z_n))^2, 2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n)) \\
&\quad + 2\text{Cov}((\bar{c}(z_{n+1} - z_n))^2, (\varepsilon_{n+1}^* - \varepsilon_n^*)^2) \\
&\quad + 2\text{Cov}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n), (\varepsilon_{n+1}^* - \varepsilon_n^*)^2).
\end{aligned}$$

but for similar reasons as above, all covariances are zero. Note also that $\text{Var}(2(\varepsilon_{n+1}^* - \varepsilon_n^*)\bar{c}(z_{n+1} - z_n)) = 4\text{Var}(\varepsilon_{n+1}^* - \varepsilon_n^*)\text{Var}(\bar{c}(z_{n+1} - z_n)) = 8\text{Var}(\varepsilon^*)\text{Var}(\bar{c}(z_{n+1} - z_n))$. Also note that:

$$\begin{aligned}
\text{Var}((\varepsilon_{n+1}^* - \varepsilon_n^*)^2) &= \mathbb{E}[(\varepsilon_{n+1}^* - \varepsilon_n^*)^4] - \mathbb{E}[(\varepsilon_{n+1}^* - \varepsilon_n^*)^2]^2 \\
&= 2\mathbb{E}[(\varepsilon^*)^4] + 6\mathbb{E}[(\varepsilon^*)^2]^2 - (2\mathbb{E}[(\varepsilon^*)^2])^2 \\
&= 2(\mathbb{E}[(\varepsilon^*)^4] + \mathbb{E}[(\varepsilon^*)^2]^2) = 2\text{Var}(\varepsilon^*)^2(\text{Kurt}(\varepsilon^*) + 1).
\end{aligned}$$

Thus, putting everything together:

$$\tilde{\beta}_2 = \frac{\bar{c}^4 \text{Cov}((z_{n+1} - z_n)^2, (z_n - z_{n-1})^2) - 4\bar{c}^2 \text{Var}(\varepsilon^*) \text{Cov}((z_{n+1} - z_n), (z_n - z_{n-1})) + \text{Var}(\varepsilon^*)^2[\text{Kurt}(\varepsilon^*) - 1]}{\text{Var}((\bar{c}(z_{n+1} - z_n))^2) + 8\text{Var}(\varepsilon^*)\text{Var}(\bar{c}(z_{n+1} - z_n)) + 2\text{Var}(\varepsilon^*)^2(\text{Kurt}(\varepsilon^*) + 1)}.$$

CALIBRATING THE PARAMETERS FROM THE FORMULAS. I take κ as something we calibrate exogenously from the frequency, ρ as fixed, and η as something we calibrate externally.

Our remaining observables are $\text{Var}(\Delta p)$, $\text{Kurt}(\Delta p)$, β_1 and β_2 . First, for $\text{Var}(\Delta p)$ we have:

$$\text{Var}(\Delta p) = \bar{c}^2 \left(\frac{\sigma^2}{\kappa + \eta} \right) + 2\text{Var}(\varepsilon)$$

and from $\text{Cov}(\Delta p_{t+\tau}, \Delta p_t) = -(\frac{\sigma^2}{2\eta} c^2 (\frac{\eta}{\eta+\kappa})^2 + \text{Var}(\varepsilon^*)) = -(c^2 \frac{\sigma^2}{2} \frac{\eta}{(\kappa+\eta)^2} + \text{Var}(\varepsilon^*))$ so from this, we have:

$$\begin{aligned} \text{Var}(\Delta p)(1 + 2\beta_1) &= \bar{c}^2 \left(\frac{\sigma^2}{\kappa + \eta} \right) - c^2 \frac{\sigma^2}{2} \frac{\eta}{(\kappa + \eta)^2} \\ \text{Var}(\Delta p)(1 + 2\beta_1) &= \bar{c}^2 \sigma^2 \frac{\kappa}{(\kappa + \eta)^2} \end{aligned}$$

which directly identifies σ^2 . Given that, we can then identify $\text{Var}(\varepsilon^*)$ directly. We then have one parameter ($\text{Kurt}(\varepsilon^*)$) to match both β_2 and $\text{Kurt}(\Delta p)$. I match $\tilde{\beta}_2$ and then see how far off I am in the kurtosis. For the elements of $\tilde{\beta}_2$, we have the following closed-form expressions:

$$\begin{aligned} \text{Cov}((z_{t+\tau_{n+1}} - z_{t+\tau_n})^2, (z_{t+\tau_n} - z_{t+\tau_{n-1}})^2) &= 2\text{Var}(z)^2 \mathbb{E}[(1 - e^{-\eta\Delta_{n+1}})^2]^2 \\ &= 2\text{Var}(z)^2 \left(\frac{2\eta^2}{(\kappa + \eta)(\kappa + 2\eta)} \right)^2 \end{aligned}$$

$$\begin{aligned} \text{Var}((z_{t+\tau_{n+1}} - z_{t+\tau_n})^2) &= 8\text{Var}(z)^2 \mathbb{E}[(1 - e^{-\eta\Delta_{n+1}})^2] + 4\text{Var}(z)^2 \text{Var}((1 - e^{-\eta\Delta_{n+1}})) \\ &= 4\text{Var}(z)^2 \left(\frac{(5\kappa + 4\eta)}{(\kappa + \eta)^2 (\kappa + 2\eta)} \eta^2 \right) \\ \text{Cov}((z_{t+\tau_{n+1}} - z_{t+\tau_n}), (z_{t+\tau_n} - z_{t+\tau_{n-1}})) &= -\text{Var}(z) (\eta / (\eta + \kappa))^2 \\ \text{Var}((z_{t+\tau_{n+1}} - z_{t+\tau_n})) &= 2\text{Var}(z) \left(\frac{\eta}{\kappa + \eta} \right) \end{aligned}$$

so using that we can recover $\text{Kurt}(\varepsilon)$ from $\tilde{\beta}_2$ as:

$$\begin{aligned} \beta_2 [\text{Var}(\bar{c}(z_{t+\tau_{n+1}} - z_{t+\tau_n}))^2 + 8\text{Var}(\varepsilon^*)\text{Var}(\bar{c}(z_{t+\tau_{n+1}} - z_{t+\tau_n})) \\ + 2\text{Var}(\varepsilon^*)^2 (\text{Kurt}(\varepsilon) + 1)] = \\ \bar{c}^4 \text{Cov}((z_{t+\tau_{n+1}} - z_{t+\tau_n})^2, (z_{t+\tau_n} - z_{t+\tau_{n-1}})^2) \\ - 4\bar{c}^2 \text{Var}(\varepsilon^*) \text{Cov}((z_{t+\tau_{n+1}} - z_{t+\tau_n}), (z_{t+\tau_n} - z_{t+\tau_{n-1}})) + \text{Var}(\varepsilon^*)^2 [\text{Kurt}(\varepsilon^*) - 1] \end{aligned}$$

so

$$\text{Kurt}(\varepsilon^*) = (\text{Num})/(\text{Var}(\varepsilon^*)^2(2\beta_2 - 1)).$$

where

$$\begin{aligned} \text{Num} = & -\beta_2[\text{Var}(\bar{c}(z_{t+\tau_{n+1}} - z_{t+\tau_n}))^2] + 8\text{Var}(\varepsilon^*)\text{Var}(\bar{c}(z_{t+\tau_{n+1}} - z_{t+\tau_n})) + 2\text{Var}(\varepsilon^*)^2] \\ & + \bar{c}^4 \text{Cov}((z_{t+\tau_{n+1}} - z_{t+\tau_n})^2, (z_{t+\tau_n} - z_{t+\tau_{n-1}})^2) \\ & - 4\bar{c}^2 \text{Var}(\varepsilon^*) \text{Cov}((z_{t+\tau_{n+1}} - z_{t+\tau_n}), (z_{t+\tau_n} - z_{t+\tau_{n-1}})) - \text{Var}(\varepsilon^*)^2. \end{aligned}$$

OA.D Microfoundations and Approximations

The analysis in the main text uses several simplifying assumptions and approximations, such as quadratic profits, continuous time, and computing the IRF based on linear formulas. The underlying microfoundations follow from Alvarez and Lippi (2014). The specification for the control costs, which allows us to write that problem as a function of a single state, is novel to this paper. I also present the underlying discrete time approximation to the continuous time specification used in the paper.

OA.D.1 General Equilibrium Setup

Time is continuous and indexed by $t \in [0, +\infty)$. There is a continuum of households with discount rate ρ and utility $\int_0^\infty e^{-\rho t} [\log(C(t)) - \alpha L(t) + \log(\frac{M(t)}{P(t)})] dt$, where $C(t)$ is a consumption composite, $L(t)$ is labor, $M(t)$ is the nominal money stock and $P(t)$ is the aggregate price level. There is a continuum of firms $k \in [0, 1]$. Each firm sells $n \geq 1$ different products, indexed by i . The consumption composite is given by $C(t) = \left(\int_0^1 \sum_{i=1}^n (Z_{k,i}(t) c_{k,i}(t))^{\frac{\xi-1}{\xi}} dk \right)^{\frac{\xi}{\xi-1}}$ where $Z_{k,i}(t)$ are preference shocks as standard in the literature. Households maximize utility by choosing $c_{k,i}(t)$, $L(t)$, $M(t)$ subject to a budget constraint:

$$M(0) + \int_0^\infty Q(t) [\tilde{\Pi}(t) + (1 + \tau_L)W(t)L(t) - R(t)M(t) - \int_0^1 \sum_{i=1}^n \tilde{p}_{k,i}(t) c_{k,i}(t) dk] dt = 0,$$

where $Q(t) = e^{-\int_0^t R(s) ds}$, $R(t)$ is the nominal interest rate, $\tilde{\Pi}(t)$ is aggregate profits, τ_L is a labor subsidy, $W(t)$ is the nominal wage, and $\tilde{p}_{k,i}(t)$ are the prices of good i sold by firm k . Note that due to the homotheticity of the aggregator, $\int_0^1 \sum_{i=1}^n \tilde{p}_{k,i}(t) c_{k,i}(t) dk = P(t)C(t)$. Letting λ_0 be the Lagrange multiplier of the aggregate budget constraint, the FOCs for the

household imply:

$$\begin{aligned}
\frac{e^{-\rho t}}{C(t)} &= \lambda_0 Q(t) P(t) \\
\frac{e^{-\rho t}}{M(t)} &= \lambda_0 Q(t) R(t) \\
\alpha e^{-\rho t} &= \lambda_0 (1 + \tau_L) Q(t) W(t) \\
e^{-\rho t} \frac{1}{C(t)} C(t)^{1/\xi} c_{k,i}^{-1/\xi} (Z_{k,i}(t))^{1-1/\xi} &= \lambda_0 Q(t) \tilde{p}_{k,i}(t).
\end{aligned}$$

Assume that money follows $M(t) = M(0) \exp(\mu t)$. This implies that for all t , $R(t) = \rho + \mu$, $W(t) = \frac{\alpha}{1+\tau_L} (r + \mu) M(t)$ and $\lambda_0 = ((\rho + \mu) M(0))^{-1}$. From the FOCs for consumption and labor we have $C(t) = \frac{1+\tau_L}{\alpha} \frac{W(t)}{P(t)}$. Using the FOCs for $c_{k,i}$, $C(t)$ and the last equation we get $c_{k,i}(t) = (C(t))^{1-\xi} (Z_{k,i}(t))^{\xi-1} \left(\frac{\alpha}{1+\tau_L} \frac{\tilde{p}_{k,i}(t)}{W(t)} \right)^{-\xi}$. When considering the effect of a monetary shock, I assume $\mu = 0$, and that initially $M(0)/P(0)$ (where $P(0)$ is consistent with the invariant distribution of price gaps) is at steady state. I then consider a perturbed initial condition $M(0)e^\delta$ where δ is small.

The production function for good i in firm k at time t is linear in labor with productivity $1/Z_{k,i}(t)$. Thus, $Z_{k,i}(t)$ indexes shocks to both preferences and technology. This assumption is standard in the literature, and it allows me to reduce the firm's problem to a single state variable when $Z_{k,i}$ follows a random walk. The nominal marginal cost is thus $W(t)Z_{k,i}(t)$. The nominal profit for a firm k selling product i at price $\tilde{p}_{k,i}$ is $(C(t))^{1-\xi} (Z_{k,i}(t))^{\xi-1} \left(\frac{\alpha}{1+\tau_L} \frac{\tilde{p}_{k,i}(t)}{W(t)} \right)^{-\xi} [\tilde{p}_{k,i}(t) - W(t)Z_{k,i}(t)]$. Rearranging yields:

$$W(t)(C(t))^{1-\xi} \left(\frac{\alpha}{1+\tau_L} \frac{\tilde{p}_{k,i}(t)}{Z_{k,i}(t)W(t)} \right)^{-\xi} \left[\frac{\tilde{p}_{k,i}(t)}{W(t)Z_{k,i}(t)} - 1 \right].$$

Given the CES structure, the static profit-maximizing price is $p_{k,i}^*(t) = \frac{\xi}{\xi-1} Z_{k,i}(t) W(t)$. Letting $x_{k,i}(t) = \log \frac{\tilde{p}_{k,i}(t)}{p_{k,i}^*(t)}$ be the price gap, the profit of a given product sold by a given firm is $W(t)(C(t))^{1-\xi} \Pi(x_{k,i})$ where:

$$\Pi(x) = \left(\frac{\alpha}{1+\tau_L} \frac{\xi}{\xi-1} \right)^{-\xi} e^{-\xi x} \left(\frac{\xi}{\xi-1} e^x - 1 \right).$$

Note also that $\Pi(x)/\Pi(0) = e^{-\xi x} (1 + \xi e^x - \xi)$, which I use below. Given this, the aggregate price index $P(t)$ can be entirely characterized as a function of $W(t)$ and price gaps. In

particular, the ideal price index satisfies:

$$P(t) = W(t) \frac{\xi}{\xi - 1} \left(\int_0^1 \sum_{i=1}^n e^{(1-\xi)x_{k,i}} dk \right)^{1/(1-\xi)}.$$

OA.D.1.1 Firm's Problem

The firm's problem is to choose stopping times at which prices change, $\{\tau_j\}_{j=1}^\infty$ and, at time τ_j conditional reset-gap distributions for each product, $\{f_{j,i}^*\}_{j=1}^\infty$ in order to maximize the value of its nominal profits:

$$\mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} \left(\sum_{i=1}^n W(t) (C(t))^{1-\xi} \Pi(x_i) \right) dt - \sum_{j=1}^\infty W(\tau_j) e^{-\rho \tau_j} \sum_{i=1}^n \mathcal{C}_\ell(f_{j,i}^* || \bar{f}_j; \zeta(\tau_j)) \right].$$

The functional $\mathcal{C}_\ell(f_{j,i}^* || \bar{f}_j; \zeta(\tau_j))$ encodes costs of changing prices, for example menu costs or control costs. It depends on the chosen conditional distribution for reset gaps $f_{j,i}^*$, as well as the baseline distribution $\bar{f}_j$, which is taken as given by the firm. I specify the characteristics of the baseline distribution below.

The process $\zeta(t)$ indexes randomness in price adjustment costs. I discuss some specifications below. Let $\hat{\Pi}(x_i) = \Pi(x_i)/\Pi(0)$, using that equilibrium nominal wages are constant $W(t)/\bar{W} = e^\delta$ and rescaling the cost function appropriately (and letting the rescaled function be $\mathcal{C}$), we can define the value (per product) as:

$$\mathcal{V} = \frac{1}{n} \Pi(0) \bar{W} e^\delta \bar{C}^{1-\xi} \mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} \left(\sum_{i=1}^n \left(\frac{C(t)}{\bar{C}} \right)^{1-\xi} \hat{\Pi}(x_i) \right) dt - \sum_{j=1}^\infty e^{-\rho \tau_j} \sum_{i=1}^n \mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(\tau_j)) \right],$$

where $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(\tau_j)) = \frac{1}{\Pi(0) \bar{C}^{1-\xi}} \mathcal{C}_\ell(f_{j,i}^* || \bar{f}_j; \zeta(\tau_j))$ is the rescaled version of the cost of changing prices. Letting $\mathcal{S}(C(t), x_i) = (\frac{C(t)}{\bar{C}})^{1-\xi} \hat{\Pi}(x_i) = (\frac{C(t)}{\bar{C}})^{1-\xi} e^{-\xi x} (1 + \xi e^x - \xi)$, we can expand it around $C(t) = \bar{C}$ and $x = 0$ as

$$\mathcal{S}(C(t), x_i) = 1 + (1-\xi) \left(\frac{C(t) - \bar{C}}{\bar{C}} \right) + \frac{1}{2} (1-\xi)^2 \left(\frac{C(t) - \bar{C}}{\bar{C}} \right)^2 - \frac{\xi(\xi-1)}{2} x_i^2 + o(\|C(t) - \bar{C}, x_i\|^2),$$

where I used that $\Pi'(0) = 0$, $\Pi''(0) = -\xi(\xi - 1)$ and thus $\frac{\partial^2}{\partial C(t) \partial x_i} \mathcal{S}(C(t), x_i) = 0$. Letting:

$$V = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \frac{1}{n} \sum_{i=1}^n (-Bx_i^2) dt - \sum_{j=1}^\infty e^{-\rho \tau_j} \frac{1}{n} \sum_{i=1}^n \mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(\tau_j)) \right],$$

we can write:

$$\mathcal{V} = \Pi(0) \bar{W} e^{\delta \bar{C} - \xi} V + \iota(\delta, \{C(t)\}) + \int_0^\infty e^{-\rho t} o(\|C(t) - \bar{C}, x_i\|^2) dt,$$

where $\iota(\delta, \{C(t)\})$ collects all terms from $\mathcal{S}(C(t), x_i)$ that are independent of x . Note that, for particular choices of $\mathcal{C}$ and some assumptions on the law of motion of x_i , if $B = \xi(\xi - 1)/2$, V satisfies the recursive formulas derived above, so this provides a formal justification for the quadratic profit function as a second-order approximation to the true profit function. For this approximation to be valid, deviations from the optimal price must not be too large. This can be justified by taking the variance of idiosyncratic shocks and price adjustment costs to zero at the correct rate. See Auclert et al. (2024) Online Appendix E for details.

OA.D.1.2 Specifications of Price Adjustment Cost Functional

In this section I describe several specific patterns of adjustment costs nested in the above notation. Let $\zeta(t)$ be a compound Poisson process: jumps occur with arrival rate κ , and when a jump occurs, the size of the jump is ψ , drawn from a distribution $G(\psi)$. Thus $\zeta(t) = \sum_{i=1}^{N(t)} \psi_i$ where $N(t)$ is the associated Poisson counting process.

CALVO PRICING. Calvo (1983) pricing is obtained by assuming that, if $d\zeta > 0$, $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = 0$, otherwise $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = +\infty$.

FIXED MENU COSTS. A model with fixed menu costs (Goloso and Lucas, 2007) is obtained by assuming $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = \Psi$ for all ζ .

RANDOM MENU COSTS. The random menu cost specification in Alvarez et al. (2021) is obtained by assuming that if $d\zeta > 0$, $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = \psi_j$ and otherwise $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = \Psi$.

MIXED CONTROL COST, MENU COST MODEL. The mixed model in Section OA.B.3 is obtained by assuming $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = \theta(1 - \omega)\mathcal{I}(f_{j,i}^* || \bar{f}_j) + \omega\hat{c}(\theta)$ for a fixed $\bar{f}_j$, and for all $\zeta(t)$.

MIXED CONTROL COST, RANDOM MENU COST MODEL. The mixed model in Section 5.2 is obtained by assuming $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = \theta \mathcal{I}(f_{j,i}^* || \bar{f}_j) + \psi_j$ for a fixed $\bar{f}_j$ if $d\zeta(t) > 0$, and $\mathcal{C}(f_{j,i}^* || \bar{f}_j; \zeta(t)) = \theta \mathcal{I}(f_{j,i}^* || \bar{f}_j) + \Psi$ otherwise.

OA.D.1.3 Specification of the Baseline Distribution

Besides assuming that x_i follow random walks, the other important assumption that allows us to rewrite the problem as a function of a single state variable is to appropriately scale $\bar{f}_j$ so that everything only depends on the price gap. In order to see why, from the above problem, assume that the firm's profits only depend on $x = p - z$. Then I assume that the reference distribution is indexed to z as $\bar{f}_j(p|z) = \bar{f}(p - z) = \bar{f}(x)$. For the case in which z is mean-reverting and there are control costs (Section 5.7), I assume that $\bar{f}_j(p|z) = \bar{f}(p - \bar{c}z)$, where $\bar{c}z$ is the optimal reset point in the pure Calvo model.

OA.D.2 Continuous Time as a Limit of Discrete Time

I write the model in discrete time, and then I will take the limit to get the continuous time formulation. I consider the model without random menu costs, but the derivation can be extended to that case.⁸ Periods have length Δ . Assume each time the firm changes prices, it solves a control cost problem with a cost functional $\mathcal{C}(F)$, with $\mathcal{C}(F) \geq c > 0$ for all F . Let x be the price gap, and $\pi(x)\Delta$ be profits per period. Assume $\pi(x)$ is bounded.⁹ Then the problem of the firm can be written recursively as:

$$\begin{aligned} V(x) &= \max \{ \pi(x)\Delta + e^{-\rho\Delta} \mathbb{E}[V(x + \Delta W)], V^* \} \\ \text{with } V^* &= \max_F \int (\pi(x)\Delta + e^{-\rho\Delta} \mathbb{E}[V(x + \Delta W)]) dF(x) - \mathcal{C}(F), \end{aligned}$$

where ΔW is the typical binomial step that converges to a Brownian motion increment.¹⁰ Taking the limit as $\Delta \rightarrow 0$:

⁸With random menu (or control) costs, one needs to assume that in each period, the firm can adjust prices with probability 1 by using a (high) cost functional $\bar{\mathcal{C}}$. If it chooses not to, it enters a lottery: with probability $1 - e^{-\kappa\Delta}$, it draws a cost functional from G , and then decides whether to adjust or not at that cost. With the complementary probability, the firm gets no draw and thus has to keep prices fixed.

⁹This is required for technical reasons. An economic justification is that maximum profits are naturally bounded, and that losses cannot be unboundedly large because the firm would just shut down.

¹⁰In particular, ΔW can take values $\sigma\sqrt{\Delta}$ or $-\sigma\sqrt{\Delta}$ with probabilities $p = \frac{1}{2}$ and $1 - p$ respectively.

$$\begin{aligned}
V(x) &= \max \{ \pi(x)dt + (1 - \rho dt)\mathbb{E}[V(x) + dV], V^* \} \\
0 &= \max \left\{ \left(\pi(x) + \frac{\sigma^2}{2}V_{xx} - \rho V \right)dt, V^* - V(x) \right\},
\end{aligned}$$

and we can drop the dt since at the solution the maximizer is zero, so the other argument must be weakly negative, and thus the dt does not matter. Notice that once we take the limit for V^* we get:

$$\begin{aligned}
\lim_{\Delta \rightarrow 0} V^* &= \max_F \int (\pi(x)\Delta + e^{-\rho\Delta}\mathbb{E}[V(x + \Delta W)])dF(x) - \mathcal{C}(F) \\
&= \max_F \int V(x)dF - \mathcal{C}(F),
\end{aligned}$$

where we can interchange the max and the limit because given my assumptions, $(\pi(x)\Delta + e^{-\rho\Delta}\mathbb{E}[V(x + \Delta W)])dF$ converges uniformly, i.e., $\lim_{\Delta \rightarrow 0} \sup_x |\pi(x)\Delta + e^{-\rho\Delta}[pV(x + \sqrt{\Delta}\sigma) + (1 - p)V(x - \sqrt{\Delta}\sigma)] - V(x)| = 0$ ¹¹

OA.E Numerical Solution Procedure

In this section, I provide details about the numerical solution procedure. I use numerical solutions for the $n = 1$ column of Table 2, and for the menu cost model with mean-reverting shocks in Appendix B.1.1. In one case I solve an HJBQVI in one variable, and in the other, one in two variables. I treat those cases separately.

The reset-value notation follows the convention above: $V^*(\psi)$ is net of control costs and menu cost ψ , while V^{**} is gross of those costs. In the algorithms, $v^{*,n}$ denotes the zero-menu-cost net value $V^*(0)$ at iteration n ; subtracting Ψ therefore gives the full-cost reset value. In the two-state model the same convention applies conditional on z , and control costs are zero.

¹¹In particular, we assume $\pi(x)$ is bounded, which implies V is. This strictly rules out several basic functions, but using an arbitrarily low lower bound we can get a good approximate solution to this. Furthermore, in the numerical implementation we restrict the state space to live in a closed interval on the real line, which essentially serves the same purpose as bounding π by its minimum value attained in the interval.

OA.E.1 One Variable

VALUE FUNCTION. Take the state to be x , the price gap. The HJBQVI is:

$$0 = \min\{\rho V - (\pi(x) + \frac{\sigma^2}{2}V_{xx} + \kappa \int_0^\Psi \max\{V^*(\psi) - V, 0\}dG(\psi)), V - V^*(\Psi)\}.$$

I solve this using value function iteration on a grid, using a semi-implicit scheme with finite differences. I take an equally spaced grid that ranges from $x_1 = -1.5$ to $x_{N_x} = 1.5$ with increments of $\Delta_x = 0.01$. Let π be a vector that contains the profit function evaluated at each gridpoint. Define also a time step $\Delta_t = 1/360$ (so each step is a day). To deal with the intermittent adjustment, I use an adaptation of the procedure explained in Moll (2017), which is based on Huang and Pang (1998). These essentially show that each step of the value function iteration reduces to solving a Linear Complementarity Problem (LCP), i.e., for given $\mathbf{B}$ and q , find z that satisfies:

$$z'(\mathbf{B}z + q) = 0 \tag{OA.E.1}$$

$$z \geq 0 \tag{OA.E.2}$$

$$\mathbf{B}z + q \geq 0, \tag{OA.E.3}$$

where $\mathbf{B}$ is an $N_z \times N_z$ matrix, z and q are $N_z \times 1$ vectors, and the inequalities mean that all elements of the corresponding vector are nonnegative. Given this, Algorithm OA.1 explains how to solve for the value function.

Algorithm OA.1. Value Function Iteration: One State. Before starting the main loop, precompute the $N_x \times N_x$ matrices $\mathbf{A}$ and $\mathbf{B} = (\rho + 1/\Delta_t)\mathbf{I} - \mathbf{A}$. The matrix $\mathbf{A}$ corresponds to an appropriate finite difference approximation of the operator $\mathcal{L}[V] = \frac{\sigma^2}{2}V_{xx}$, in this case for $i = 1, \dots, N_x$, $A_{i,i-1} = A_{i,i+1} = \frac{\sigma^2}{2(\Delta_x)^2}$, $A_{i,i} = -\sigma^2/(\Delta_x)^2$, and the rest of the elements are zero. Initialize the algorithm with a guess v^0 . Then, for any given guess v^n :

- i) Compute $v^{\max,n} = \max v^n$ and $v^{*,n} = v^{\max,n} + \theta \log(\int \exp((v^n(x) - v^{\max,n})/\theta) \bar{f}(x) dx)$
- ii) Compute the expected value of adjusting for each v_i^n , $E_i^n = \int_0^\Psi \max\{v^{*,n} - \psi - v_i^n, 0\}dG(\psi)$.
- iii) Construct $S^n = v^{*,n} - \Psi$ and $q^n = \mathbf{B}S^n - v^n/\Delta_t - (\pi + \kappa E^n)$. Find the solution to the LCP (OA.E.1)-(OA.E.3) using an LCP solver with $\mathbf{B}$ and q^n . Denote the solution z^n .
- iv) Update the value function to $v^{n+1} = z^n + S^n$. Compute Error = $\max_i |v_i^{n+1} - v_i^n|$.

If Error is below tolerance, stop. Otherwise, return to step i).

I use a tolerance of 10^{-8} . I compute the integrals using Gauss–Legendre quadrature: I compute the GL nodes and weights in the interval $[x_1, x_{N_x}]$. I use 200 nodes. Letting x_i^{GL}, w_i^{GL} be the nodes and weights respectively, I evaluate the integral as $\int \exp((v^n(x) - v^{\max, n})/\theta) \bar{f}(x) dx \approx \sum_{i=1}^{N_{GL}} w_i^{GL} \exp((v^n(x_i^{GL}) - v^{\max, n})/\theta) \bar{f}(x_i^{GL})$, where $v^n(x_i^{GL})$ is computed from the grid via linear interpolation. For the random menu costs, I use a discrete distribution, so the integral is a sum weighted by the corresponding probabilities. This algorithm yields a vector v which approximates the value function on the gridpoints. As a byproduct, at the last step I also store $v^{\max}, v^*$ and $\mathbf{A}$.

HAZARD FUNCTION. With the objects mentioned above, we can directly compute the hazard function at gridpoint i as $\Lambda(x_i) = \Lambda_i = \kappa \sum_{j=1}^{N_\psi} \mathbf{1}_{v^* - \psi_j - V_i > 0} \text{Prob}(\psi = \psi_j)$. I then account for the strict inaction threshold by setting $\Lambda_i = 1/\Delta_t$ for all i such that $z_i = 0$ in the LCP solution at the final iteration.

RESET GAP DISTRIBUTION. For each gridpoint, I construct the unnormalized density $\tilde{f}$ as $\tilde{f}_i = \bar{f}(x_i) \exp((v_i - v^{\max})/\theta)$ and then the normalized density $f_i = \tilde{f}_i / (\sum_{i=1}^{N_x} \tilde{f}_i \Delta_x)$. I then compute a probability mass function of landing in each gridpoint (which is what I use computationally) as $dF_i = \int_{x_i - \Delta_x/2}^{x_i + \Delta_x/2} f(x) dx$ where $f(x)$ is constructed from $\tilde{f}_i$ via linear interpolation, and the integral is computed via a trapezoidal approximation.

STATIONARY DISTRIBUTION, PRICE CHANGES, AND REGRESSION COEFFICIENTS. Given that, I adapt the algorithm from Moll (2017) as follows: starting a period, there are two subperiods or “steps”:

1. Apply the diffusion operator A^T .
2. The decision of whether to change prices is made: for each x , a fraction $\Lambda(x)\Delta_t$ adjusts, and the rest remains at its current price. Out of the fraction that adjusts, they go to x' , where x' is drawn from $F(x)$.

In matrix terms, let g^n be the distribution at step n . Then:

$$\begin{aligned} g^{n+1/2} &= g^n + A^T \Delta_t g^{n+1/2} \\ g^{n+1} &= C g^{n+1/2}, \end{aligned}$$

with $C = I - \Delta_t \text{diag}(\Lambda) + \Delta_t dF \Lambda'$, where Λ is a vector that evaluates Λ in the gridpoints, and dF is a vector with elements equal to the probability of x falling in that gridpoint under the stochastic choice, constructed above. Notice that since dF sums up to 1, then the columns of C all sum up to 1, and, provided that $0 \leq \Lambda_i \Delta_t \leq 1$ for every gridpoint, its entries are nonnegative. Under this time-step condition it is a proper transition matrix. We can do an implicit step in the first part, so $g^{n+1/2} = (I - A^T \Delta_t)^{-1} g^n$ and thus $g^{n+1} = C(I - A^T \Delta_t)^{-1} g^n$. We can find the stationary distribution by iterating this update. Given any distribution g , we can compute several moments of interest.

1. The distribution of x conditional on a price change at time t is $\bar{g} = \Lambda \cdot g^{pre} / (\Lambda' g^{pre})$ where $g^{pre} = (I - A^T \Delta_t)^{-1} g$ is the distribution after idiosyncratic shocks hit but before any price adjustment was made, $\Lambda \cdot g^{pre}$ is the elementwise product of Λ and g^{pre} .
2. The frequency of price changes at each Δ_t is $f = \Delta_t \Lambda' g^{pre}$, and thus the number of price changes per period is $N = f / \Delta_t$.
3. Letting f be the density of x conditional on a price change, the density of the distribution of price changes is $f(\Delta p) = \sum_{x \in x} \bar{g}_x \times f_1(x + \Delta p)$ where f_1 is the density of the reset x conditional on reset. Basically, for each x we ask: what is the “probability” density that the price increase is Δ ? And then we integrate (sum) over all x to obtain the total density.
4. Given that, we can just compute the share of price changes that are increases/decreases and then obtain frequencies of price increases/decreases.
5. We can also compute the average size of price increases, decreases, etc., all from $f(\Delta p)$.
6. For predictability coefficients, since we use a model with no inflation trend, I compute $\mathbb{E}[x^*] = \sum_{i=1}^{N_x} x_i dF_i$, $\text{Var}(x^*) = \sum_{i=1}^{N_x} (x_i - \mathbb{E}[x^*])^2 dF_i$, $\text{Kurt}(x^*) = (\sum_{i=1}^{N_x} (x_i - \mathbb{E}[x^*])^4 dF_i) / \text{Var}(x^*)^2$. Given this, I use the formulas in Proposition 10 directly to compute β_1 and β_2 .

OA.E.2 Two Variables

I write the theory in terms of $q = \log(P/(W\xi/(\xi - 1)))$ and z , thus we have:

$$\begin{aligned} dq &= -\mu dt \\ dz &= -\eta(z - \bar{z})dt + \sigma dW, \end{aligned}$$

and the HJBQVI for a general random menu cost model is

$$0 = \max \left\{ F(q, z) - \mu V_q - \eta(z - \bar{z})V_z + \frac{\sigma^2}{2}(V_{zz}) + \kappa \int \max \{ V^*(z, 0) - V(q, z) - \tilde{\psi}, 0 \} dG(\tilde{\psi}) - \rho V, \right. \\ \left. V^*(z, \Psi) - V(q, z) \right\}. \quad (\text{OA.E.4})$$

For numerical implementation, we create a two-dimensional grid on $X \times Z$ where $x = q - z$. Importantly, we make the step size the same for both dimensions, to avoid any interpolation. My grid for x goes from $x_1 = -0.5$ to $x_{N_x} = 0.5$ in increments of $\Delta_x = 0.02$. The time step is the same as before. My grid for Z runs from $z_1 = \text{floor}(-3\sqrt{\sigma^2/(2\eta)}/\Delta_x)\Delta_x$ to $z_{N_z} = -z_1$, i.e., I essentially cover three standard deviations of the stationary distribution of z on each side. I numerically solve the problem with states $X \times Z$, in order to keep a smaller grid in X . This way, I approximate the derivatives as:

- For V_p , we simply use an upwind scheme. If $-\mu < 0$, use the backward difference, $V_p = \frac{v_{i,j} - v_{i-1,j}}{\Delta x}$, if $-\mu > 0$ use the forward difference: $V_p = \frac{v_{i+1,j} - v_{i,j}}{\Delta x}$.
- For the derivatives with z , note that fixing q means moving x in the $X \times Z$ grid. Thus, we can use an upwind scheme for $-\eta(z - \bar{z})V_z$ as follows. If $-\eta(z - \bar{z}) < 0$, use the backward difference $V_z = \frac{v_{i,j} - v_{i+1,j-1}}{\Delta z}$, otherwise use the forward difference $V_z = \frac{v_{i-1,j+1} - v_{i,j}}{\Delta z}$ where note that we are moving in the i direction with the opposite sign. For the second derivative, we can use the central difference $V_{zz} = \frac{v_{i-1,j+1} - 2v_{i,j} + v_{i+1,j-1}}{(\Delta z)^2}$.
- For boundary conditions, I use a reflecting boundary for z (so $V_z = 0$ for all x when z is at the boundary) and use that the grid is large enough for x to also impose $V_x = 0$ at the boundaries of x , since the idea is that we are in the adjustment region with probability 1, and thus V is invariant to x in this region since we adjust anyway. In particular, this implies the following adjustments when constructing the first and second derivatives. If $i = 1$ or I then $V_x = 0$, which we account for by putting zeros in the corresponding row. For derivatives with respect to z : for the first derivative use that $v_{0,j} = v_{1,j}$ and $v_{I+1,j} = v_{I,j}$ and replace that in the proper upwinding scheme. The same goes for the second derivatives.
- If $j = 1$, then $V_{zz} = \frac{v_{i-1,j+1} - v_{i,j}}{(\Delta z)^2}$, if $j = J$ then $V_{zz} = \frac{-v_{i,j} + v_{i+1,j-1}}{(\Delta z)^2}$. Note that this last part does not work if we are at the boundary for both x and z . In that case, since $v_{i-1,j+1} = v_{i,j+1}$ we use $V_{zz} = \frac{v_{i,j+1} - v_{i,j}}{(\Delta z)^2}$ or $V_{zz} = \frac{-v_{i,j} + v_{i,j-1}}{(\Delta z)^2}$ respectively.

It can be shown that this scheme produces a matrix $\mathbf{A}$ that has negative entries along the diagonal, and weakly positive entries outside of it, as well as rows that sum up to 0. Note that this scheme is essentially replicating what a standard scheme would do with (OA.E.4) if the states were (q, z) , but that representation does not have cross partials. Thus, using the simple structure of this problem, I can sidestep dealing with arbitrary cross-partial, which would introduce additional complications.

From the above discussion we can construct an $\mathbf{A}$ matrix ($N_x N_z \times N_x N_z$) that properly approximates $\mathcal{L}[V] = -\mu V_q - \eta(z - \bar{z})V_z + \frac{\sigma^2}{2}(V_{zz})$ on a joint grid for X and Z , where now v stacks all the values for $V(x, z)$, in the following order: the vector $V(x, z_1)$ (for all values of x) comes first, then $V(x, z_2)$, and so on. Given that, the algorithm has the same structure as Algorithm OA.1, with a few changes to accommodate some new objects.

Algorithm OA.2. Value Function Iteration: Two States. *Before starting the main loop, precompute the $N_x N_z \times N_x N_z$ matrices $\mathbf{A}$ and $\mathbf{B} = (\rho + 1/\Delta_t)\mathbf{I} - \mathbf{A}$. The matrix $\mathbf{A}$ corresponds to an appropriate finite difference approximation of the operator $\mathcal{L}[V] = -\mu V_q - \eta(z - \bar{z})V_z + \frac{\sigma^2}{2}(V_{zz})$. Initialize the algorithm with a guess v^0 . Then, for any given guess v^n :*

- i) For each value of z_j , compute $v_j^{*,n} = \max_x v_j^n(x)$.*
- ii) Compute the expected value of adjusting for each $v_{i,j}^n$, $E_{i,j}^n = \int_0^\Psi \max\{v_j^{*,n} - \psi - v_{i,j}^n, 0\}dG(\psi)$.*
- iii) Stacking v^n as explained above, construct $S^n = \text{Kron}(v^{*,n}, \mathbf{1}_{N_x \times 1}) - \Psi$ and $q^n = \mathbf{B}S^n - v^n/\Delta_t - (\pi + \kappa E^n)$. Find the solution to the LCP (OA.E.1)-(OA.E.3) using an LCP solver with $\mathbf{B}$ and q^n . Denote the solution z^n .*
- iv) Update the value function to $v^{n+1} = z^n + S^n$. Compute Error = $\max_i |v_i^{n+1} - v_i^n|$. If Error is below tolerance, stop. Otherwise, return to step i).*

Algorithm OA.2 explains how to solve for v in the two-state case. The tolerance is set to 10^{-5} . I replace the computation of v^* with a simple maximization, given that there is no control cost, and the only remaining difference is how to construct the matrices $\mathbf{A}$ and $\mathbf{B}$. The solution yields a value function v evaluated at the gridpoints, reset values v_j^* and the operator $\mathbf{A}$. One can obtain optimal reset points x_j^* in a variety of ways. I choose to maximize $v_j(x)$ where I interpolate between gridpoints using splines, as in each iteration of the value-function solver. This yields x_j^* that are not necessarily on the grid.

HAZARD FUNCTION. It is computed exactly as above, using $\Lambda(x, z) = \Lambda_{ij} = \kappa \sum_{k=1}^{N_\psi} \mathbf{1}_{v_j^* - \psi_k - v_{i,j} > 0} \text{Prob}(\psi = \psi_k)$, and setting it to $\Lambda_{ij} = 1/\Delta_t$ in the adjustment region. This means that the probability

of changing prices in an interval Δ_t is $\Lambda(x, z)\Delta t$. This object is a vector, stacked across states in the same way as before.

RESET POINT. I create a matrix $dF_{i,j}$ that approximates the probability of landing at i after resetting with productivity j . Given that the reset point is not necessarily on the grid, I simply divide the mass linearly between the two nearest gridpoints, that is, if $x_{i,j} < x_j^* < x_{i+1,j}$ then $dF_{i+1,j} = \frac{x_j^* - x_{i,j}}{x_{i+1,j} - x_{i,j}}$, $dF_{i,j} = 1 - dF_{i+1,j}$, and $dF_{k,j} = 0$ for all other k .

STATIONARY DISTRIBUTION, PRICE CHANGES, AND REGRESSION COEFFICIENTS. I use a procedure similar to the one above, but the operator that accounts for jumps must be modified. In particular, define $\mathbf{\Lambda} = \text{diag}(\text{vec}(\Lambda))$, a diagonal matrix that has the vectorized version of Λ along the diagonal, and $d\mathbf{F} = \text{diag}(\text{vec}(dF))$ as the analogous diagonal matrix for the reset points. Define also $\mathbf{\Lambda}_2 = \text{diag}(\mathbf{1}_{N_x \times 1} \Lambda_{\bullet,1}, \dots, \mathbf{1}_{N_x \times 1} \Lambda_{\bullet,N_z})$, a block diagonal matrix where block j is given by the row vector $\Lambda_{\bullet,j}$ stacked vertically N_x times. Now the C matrix is given by:

$$C = I - \Delta_t \mathbf{\Lambda} + \Delta_t (d\mathbf{F} \mathbf{\Lambda}_2).$$

Given this, we can find the stationary distribution by iterating on $g^{n+1} = C(I - A^T \Delta_t)^{-1} g_n$ from a g_0 that sums up to one, exactly as before. Given the stationary distribution, we can compute the following moments of interest:

1. The distribution of x conditional on a price change at time t is $\bar{g} = \Lambda * g^{pre} / (\Lambda' g^{pre})$ where $g^{pre} = (I - A^T \Delta_t)^{-1} g$ is the distribution after idiosyncratic shocks hit but before any price adjustment was made, $\Lambda * g^{pre}$ is the elementwise product of Λ and g^{pre} . Also let $\bar{\mathbf{g}}$ denote the matrix version of $\bar{g}$, i.e., we order rows and columns to correspond to values of x and z .
2. The frequency of price changes at each Δ_t is $f = \Delta_t \Lambda' g^{pre}$, and thus the number of price changes per period is $N = f / \Delta_t$.
3. In order to compute the density of price changes, note that $\Delta p = x^*(z) - x$. For each fixed z , I compute a $2N_x - 1$ vector of the joint probability of $z = z_j$ and a price change of size Δp_j occurring as:

$$\text{Prob}(\Delta p = \Delta p_i, z = z_j) = (dF_{\bullet,j} * \text{flip}(\bar{\mathbf{g}}_{\bullet,j}))[n],$$

where $(dF_{\bullet,j} * \text{flip}(\bar{\mathbf{g}}_{\bullet,j}))[n]$ is the discrete convolution between vectors $dF_{\bullet,j}$ and $\text{flip}(\bar{\mathbf{g}}_{\bullet,j})$ (where $\text{flip}(\bar{\mathbf{g}}_{\bullet,j})$ means that we take $\bar{\mathbf{g}}_{\bullet,j}$ and reverse the order of the indices). The price changes Δp_j are on a grid that spans from $2x_1$ to $2x_{N_x}$, with the same Δ_x increment. Given this, we can compute the unconditional distribution of price changes by simply summing across z_j as $\text{Prob}(\Delta p = \Delta p_i) = \sum_{j=1}^{N_z} \text{Prob}(\Delta p = \Delta p_i, z = z_j)$. Once I have that discrete distribution, computing sizes of price changes, variance, kurtosis, and other moments is straightforward.

References

- AFROUZI, H., J. P. FLYNN, AND C. YANG (2026): “What Can Measured Beliefs Tell Us About Monetary Non-Neutrality?” Working paper.
- AKERLOF, G. A. AND J. L. YELLEN (1985): “A Near-Rational Model of the Business Cycle, with Wage and Price Inertia,” *Quarterly Journal of Economics*, 100, 823–838.
- ALVAREZ, F., H. LE BIHAN, AND F. LIPPI (2016a): “The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach,” *American Economic Review*, 106, 2817–51.
- ALVAREZ, F. AND F. LIPPI (2014): “Price Setting with Menu Cost for Multi-Product Firms,” *Econometrica*, 82, 89–135.
- ALVAREZ, F., F. LIPPI, AND A. OSKOLKOV (2021): “The Macroeconomics of Sticky Prices with Generalized Hazard Functions,” *The Quarterly Journal of Economics*, 137, 989–1038.
- ALVAREZ, F., F. LIPPI, AND L. PACIELLO (2011): “Optimal Price Setting with Observation and Menu Costs,” *The Quarterly Journal of Economics*, 126, 1909–1960.
- (2016b): “Monetary Shocks in Models with Inattentive Producers,” *Review of Economic Studies*, 83, 421–459.
- AUCLERT, A., R. RIGATO, M. ROGNLIE, AND L. STRAUB (2024): “New Pricing Models, Same Old Phillips Curves?” *The Quarterly Journal of Economics*, 139, 121–186.
- BALEY, I. AND A. BLANCO (2019): “Firm Uncertainty Cycles and the Propagation of Nominal Shocks,” *American Economic Journal: Macroeconomics*, 11, 276–337.

- CALVO, G. A. (1983): “Staggered Prices in a Utility-Maximizing Framework,” *Journal of Monetary Economics*, 12, 383–398.
- COSTAIN, J. AND A. NAKOV (2015): “Precautionary Price Stickiness,” *Journal of Economic Dynamics and Control*, 58, 218–234.
- DIXIT, A. (1991): “Analytical Approximations in Models of Hysteresis,” *The Review of Economic Studies*, 58, 141–151.
- GAGLIARDONE, L., M. GERTLER, S. LENZU, AND J. TIELENS (2025): “Micro and Macro Cost-Price Dynamics in Normal Times and During Inflation Surges,” Tech. Rep. 33478, National Bureau of Economic Research.
- GOLOSOV, M. AND R. E. J. LUCAS (2007): “Menu Costs and Phillips Curves,” *Journal of Political Economy*, 115, 171–199.
- HAMILTON, J. D. (1994): *Time Series Analysis*, Princeton, NJ: Princeton University Press.
- HUANG, J. AND J.-S. PANG (1998): “Option Pricing and Linear Complementarity,” *Journal of Computational Finance*, 2, 31–60.
- KEHOE, P. J. AND V. MIDRIGAN (2015): “Prices Are Sticky After All,” *Journal of Monetary Economics*, 75, 35–53.
- LUCAS, R. E. (1972): “Expectations and the neutrality of money,” *Journal of Economic Theory*, 4, 103–124.
- MANKIW, N. G. AND R. REIS (2002): “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve,” *The Quarterly Journal of Economics*, 117, 1295–1328.
- MAĆKOWIAK, B. AND M. WIEDERHOLT (2009): “Optimal Sticky Prices under Rational Inattention,” *American Economic Review*, 99, 769–803.
- MOLL, B. (2017): “Code and Computational Resources for Heterogeneous-Agent Models,” <https://benjaminmoll.com>, accessed: 2026-02-25.
- NAKAMURA, E. AND J. STEINSSON (2008): “Five Facts About Prices: A Reevaluation of Menu Cost Models,” *Quarterly Journal of Economics*, 123, 1415–1464.
- REIS, R. (2006): “Inattentive Producers,” *Review of Economic Studies*, 73, 793–821.

- SIMS, C. A. (2003): “Implications of Rational Inattention,” *Journal of Monetary Economics*, 50, 665–690.
- STOKEY, N. L. (2009): *The Economics of Inaction: Stochastic Control Models with Fixed Costs*, Princeton, NJ: Princeton University Press.
- TAYLOR, J. B. (1979): “Staggered Wage Setting in a Macro Model,” *The American Economic Review*, 69, 108–113.
- WOODFORD, M. (2009): “Information-Constrained State-Dependent Pricing,” *Journal of Monetary Economics*, 56, 100–124.