Robust Statistics for Monetary Non-Neutrality in Menu Cost Economies

Beatrice Allamand
UC San Diego

Juan Herreño
UC San Diego and NBER

Noémie Pinardon-Touati
Columbia University

First Version: September 14th, 2026.

Abstract

The first two entries of the IRF of firm-level prices to firm-level marginal cost shocks accurately approximate the generalized Phillips curve Jacobian in a wide class of pricing models. This result is exact in Calvo and in the Gertler-Leahy menu cost model, is numerically accurate for a broad set of CalvoPlus parameterizations, including those in Golosov and Lucas (2007) and Nakamura and Steinsson (2010), and in models with micro real rigidities, shocks of different persistence, and measurement error. Our approach approximates the hazard rate of the Calvo model closest to the unknown true menu-cost Jacobian. This virtual Calvo hazard may be arbitrarily different from the frequency of price changes computed in a dataset of firm-level prices.

We thank Hassan Afrouzi, Alisdair McKay, and Matthew Rognlie for valuable comments.

1. Introduction

A large research program in monetary economics aims to infer the extent of monetary non-neutrality from unconditional moments of the distribution of price changes. The best-known example is the frequency of price changes: in the textbook Calvo (1983) model it is a sufficient statistic for monetary non-neutrality. This program extends to the more complex case of state-dependent models: there, the frequency of price change is uninformative about the degree of non-neutrality (e.g., Caplin and Spulber 1987), but—under certain conditions—the ratio of kurtosis to frequency of price changes is a sufficient statistic (Alvarez, Le Bihan, and Lippi 2016).

A key challenge in this literature is model misspecification. These unconditional moments are sufficient statistics when the model is correctly specified, but small changes in the environment can break them. Both statistics fail in the presence of micro real rigidities, a key mechanism to reconcile the modest price rigidity in micro data with large monetary non-neutrality in aggregate data (Ball and Romer 1990). The frequency is uninformative in the Rotemberg (1982) and Gertler and Leahy (2008) models, even though both share the Calvo Phillips curve. The kurtosis ratio depends on the distributional assumptions of Alvarez, Le Bihan, and Lippi (2016) and Dotsey and Wolman (2020) show that relaxing them changes the answer.

This paper takes a different route. Instead of unconditional moments of the distribution of price changes, we focus on one conditional moment: the response of a firm's price to an observed idiosyncratic shock to its cost, estimated over horizons by a local projection. Our main result is that two such local projection coefficients, at impact and one period ahead, form a *robust statistic* for the generalized Phillips curve, the first-order response of the path of inflation to the path of real marginal cost.[1] Assembling the generalized Phillips curve from these two coefficients as if the economy were Calvo recovers the true one exactly when prices are set as in Calvo (1983), as shown in Herreño, Pinardon-Touati, and Thie (2026), but also as in Rotemberg (1982) and as in Gertler and Leahy (2008), where no Calvo lottery exists. In the canonical menu cost model, the assembly is not exact but is numerically tight and extremely robust across a wide range of menu costs, free-adjustment probabilities, shock volatilities, real rigidities, and measurement error in prices.

[1] Throughout this paper, Phillips curve refers to the marginal cost-based Phillips curve. Obtaining the output-based Phillips curve additionally requires a mapping from the output gap to real marginal costs.

These conditional moments are highly robust across models because they measure the object the Phillips curve is made of. The response of the price level to an aggregate nominal cost shock is an average of the responses of individual prices to that shock, and we show that it is also the average of firms' responses to their own idiosyncratic nominal cost shock. A local projection of price changes on an observed cost shifter therefore directly measures the correct object. By contrast, an unconditional moment of the distribution of price changes is an outcome: the average of what the pricing policy produces from whatever shocks the economy contains. Recovering a response from an outcome requires a model of the policy and of the shocks.[2] In addition, only two local projection coefficients suffice because the generalized Phillips curve of menu cost models is close to having the geometric structure of the Calvo one, as shown by Auclert et al. (2024). The two entries pin down an implied survival rate of price gaps, and the Calvo structure extrapolates it with a high degree of numerical accuracy.

The approach is simple to implement. It requires a panel of firm-level prices and an observed idiosyncratic cost shifter; in our analytical derivations, the shifter is a random walk. We assume the discount factor β is known. The first step is to estimate the local projection of cumulative price changes on the shifter at impact and one period ahead, which yields two coefficients, b_0 and b_1. The second step is to compute $\phi^{\text{IRF}} = b_1/b_0 - 1$, the survival probability of prices in the Calvo economy that would generate these two coefficients, and the implied Phillips curve slope $\kappa^{\text{IRF}} = b_0(1-\beta\phi^{\text{IRF}})/\phi^{\text{IRF}}$. The third step is to proceed as if the economy were Calvo: the New Keynesian Phillips curve $\pi_t = \beta\mathbb{E}_t\pi_{t+1} + \kappa^{\text{IRF}}\, mc_t^r$, iterated forward for any path of mc_t^r, is the estimate of the generalized Phillips curve. No step requires knowing the data-generating process, and the computational cost is negligible.

We start by providing an analytical characterization of our robust statistic across models. The environment is a continuum of firms facing isoelastic demand curves, producing with labor in potentially decreasing returns to scale technologies and facing a mixture of idiosyncratic shocks. Firms minimize the profit loss from mispricing plus the cost of adjusting prices, which is zero when they draw a free adjustment opportunity and positive otherwise. Depending on the adjustment technology and on the distribution of idiosyn-

[2] For instance, such moments cannot separate strong real rigidity with large cost shocks from weak real rigidity with small ones, although the two imply different Phillips curves.

cratic shocks, the environment nests the Calvo (1983), Golosov and Lucas (2007), Gertler and Leahy (2008), and CalvoPlus (Nakamura and Steinsson 2010) models.

Our first result is an equivalence between the macro response to a macro shock and the average micro response to a micro shock. Consider the sequence-space Jacobian mapping the impulse response of nominal marginal costs to that of the price level, the nominal pass-through matrix. We show that in each of these models, under conditions on the idiosyncratic shocks that we characterize, the average response of firms' prices to their own random walk idiosyncratic cost shock—the micro IRF—identifies the response of the aggregate price level to a permanent unanticipated nominal marginal cost shock horizon by horizon—the row sums of the nominal pass-through matrix.[3] Our second result is that, although the generalized Phillips curve a priori depends on the entire nominal pass-through matrix, its first two entries suffice, either exactly or to a numerically tight approximation.

To develop intuition, we show, model by model, what the two coefficients identify. The generalized Phillips curve depends on two forces: the desired pass-through of costs into prices, $\omega \leq 1$, pinned down by real rigidities, and the decay of a cost shock in firms' price gaps, set by the adjustment technology. The impact coefficient b_0 captures both forces, while the ratio $\phi^{\text{IRF}} = b_1/b_0 - 1$ captures the decay alone. In the Calvo model, $b_0 = \omega(1-\alpha)$ and $\phi^{\text{IRF}} = \alpha$, where α is the non-adjustment probability. Because the decay is geometric, its one-period value summarizes it, and the two coefficients fully determine the Phillips curve. Importantly, both b_0 and b_1 are responses of realized prices and so reflect selection—which firms adjust after a shock and by how much—when present. In the state-dependent Gertler and Leahy (2008) model, for instance, a firm reconsiders its price only when hit by a large idiosyncratic shock, with probability $1-\eta$, and a firm that reconsiders and keeps its price had the right price on average, so that its expected price absorbs the cost shock whether or not it adjusts. The decay is geometric, at η, and the Phillips curve has the Calvo structure with η in place of α. In this case, b_0 and ϕ^{IRF} pick up $\omega(1-\eta)$ and η, although η is not the frequency of price changes, and our statistic is again exact.

In the canonical menu cost model the decay is not geometric. By the representation of Auclert et al. (2024), the pass-through matrix is a weighted average of the pass-through matrices of exactly two time-dependent models, one for the extensive margin, the move-

[3] The result is related to, but distinct from, Caballero and Engel (2007), who show that the macro response to a macro shock equals the average micro response to the same macro shock. Ours relates responses to two different shocks, which is what makes the aggregate response estimable from a firm-level regression.

ments of the adjustment bands, and one for the intensive margin, the movements of the reset price. We show that the micro IRF exactly identifies the weighted average of the two survival functions horizon by horizon, with weights equal to the shares of the impact pass-through due to each margin, and that ϕ^{IRF} is its one-period value. The Calvo assembly extrapolates this one-period value geometrically. The replacement is accurate because, as Auclert et al. (2024) document, the survival functions of the two margins deviate from a common constant in opposite directions so that their average is close to constant.

Whether the approach accurately recovers the Phillips curve of the canonical menu cost model is a quantitative question, which we answer in two steps. First, Auclert et al. (2024) show that the generalized Phillips curve of a menu cost model is well approximated by that of a virtual Calvo model with a suitably chosen hazard, which the researcher does not observe. We show that whenever this approximation is accurate, our statistic recovers the virtual hazard, and therefore the true Phillips curve, with a proportional accuracy. Second, we show numerically that the approximation is accurate across a wide range of parameterizations of the menu cost, the probability of free adjustment, and the volatility of idiosyncratic shocks, along three metrics: the impact entry of the generalized Phillips curve, entire columns of the generalized Phillips curve, and the Phillips multiplier of Barnichon and Mesters (2021), equal to the ratio of the present discounted value of the inflation response to that of the output response. Across parameterizations, the metrics computed from our assembled Phillips curve move one-for-one with the true ones and explain virtually all of their variance. Our approach does as well as a Calvo model whose hazard is chosen to best fit the true Phillips curve, which requires knowing it.

The accuracy of the approach extends beyond the canonical model. First, it accommodates decreasing returns to scale and strategic complementarities in price setting (which we model à la Kimball 1995), two forms of micro-level real rigidities, while existing sufficient statistics do not. The intuition is that firm-level pass-through already incorporates them, so that no adjustment is needed. Hence, our statistic performs as well as in the baseline model. Second, it does not require idiosyncratic cost shocks to follow a random walk, an assumption that existing sufficient statistics maintain. Third, it is unaffected by measurement error in prices. The two coefficients are conditional elasticities estimated by instrumental variables, not unconditional moments, so that noise in individual price changes does not affect them.

Related literature. This work contributes to three strands of the literature. First, we contribute to the literature that infers monetary non-neutrality from unconditional moments of the distribution of price changes. This agenda has grown alongside the evidence for state-dependent models, in which the frequency of price changes is not a sufficient statistic. In these models, the full size distribution of price changes is informative about monetary non-neutrality (Midrigan 2011; Alvarez and Lippi 2014). In particular, Alvarez, Le Bihan, and Lippi (2016) show that the ratio of kurtosis to frequency is sufficient for the cumulative response of the price level to a small monetary shock.[4] This result holds around the zero inflation steady-state and requires quadratic losses, Gaussian random-walk idiosyncratic shocks, and the absence of micro-level real rigidities so that an adjusting firm closes its price gap entirely.[5] The subsequent literature has shown that each of these assumptions is a source of misspecification. With fat-tailed shocks, Karadi and Reiff (2019) match the same frequency, size, and kurtosis of price changes with models that have very different cumulative price responses. With strategic complementarities, Klenow and Willis (2016) reproduce the same frequency, dispersion, and persistence of price changes either with no complementarity and small shocks or with a large Kimball superelasticity and large shocks. Dotsey and Wolman (2020) reach the same underidentification conclusion while targeting six moments of the distribution of price changes. Consistent with these results, Hong et al. (2023) find across producer-price sectors that only the frequency is robustly related to measured non-neutrality, and that kurtosis has little predictive power. These moments also depend on how prices are measured: Nakamura and Steinsson (2008), Eichenbaum, Jaimovich, and Rebelo (2011) and Kehoe and Midrigan (2015) show that the measured frequency of price changes is highly sensitive to the treatment of temporary sales and to the use of posted versus reference prices. Relative to this literature, we propose conditional moments, the response of prices to an observed cost shock, whose mapping to the Phillips curve is the same across time- and state-dependent models, does not depend on

[4] With the additional assumption of Golosov and Lucas (2007) preferences and given a labor supply elasticity, the response of the price level fully determines the real wage, and hence, labor supply and output.

[5] Alvarez, Lippi, and Oskolkov (2022) show that the result extends to any symmetric adjustment hazard but that the same four assumptions remain necessary. Alvarez, Lippi, and Souganidis (2023) allow for strategic complementarities, summarized by a single parameter θ, the negative of the elasticity of a firm's optimal price to the aggregate price, and find that the cumulative output response becomes the kurtosis ratio scaled by $1/\sqrt{1+\theta}$; the scaling is exact in the Calvo model and numerically close in menu cost models. However, θ is not a moment of the distribution of price changes: it must be calibrated, or estimated from the response of prices to costs and to competitors' prices, that is, from a regression of the kind we study.

the distribution of idiosyncratic shocks or on micro-level real rigidities, and is unaffected by variation in prices that is unrelated to the shock.

Second, we contribute to the literature on generalized Phillips curves in Ss models. Auclert et al. (2024) define the generalized Phillips curve as the sequence-space Jacobian from real marginal cost to inflation, and find numerically that the generalized Phillips curve of menu cost models is close to that of a Calvo model with a suitably chosen "synthetic Calvo hazard". However, operationalizing this result requires knowing the underlying model to compute the synthetic hazard that best approximates the menu cost generalized Phillips curve.[6] Our robust statistic operationalizes the Calvo approximation without knowledge of the underlying model. Our approach builds on an aggregation argument reminiscent of Caballero and Engel (2007). They show that in a stationary Ss model the impact response of the price level to a small nominal common shock is the average of firms' marginal responses to it. We show that it also equals the average response of firms' prices to their own idiosyncratic nominal cost shock, the object a cross-sectional regression identifies.

Third, we contribute to the empirical literature that estimates the pass-through of cost shocks into prices at the firm level and uses it to learn about monetary non-neutrality Gopinath and Itskhoki (2010).[7] Closest to us, Gagliardone et al. (2025) and Herreño, Pinardon-Touati, and Thie (2026) both estimate pass-through regressions to recover the slope of the marginal-cost Phillips curve in the Calvo framework. In particular, Herreño, Pinardon-Touati, and Thie (2026) show that in the Calvo model the pass-through at two horizons identifies this slope. This paper shows what the same regressions identify in state-dependent models: the Phillips curve exactly in the Gertler–Leahy model and to a tight approximation in the canonical menu cost model, so that the two-coefficient recipe applies without knowing the pricing model.[8]

[6] Alternatively, a researcher can rely on the Alvarez, Le Bihan, and Lippi (2016) sufficient statistics and the known kurtosis of the Calvo model to infer the synthetic Calvo hazard, but this approach suffers from the misspecification risk highlighted above.

[7] A large literature also studies pass-throughs outside of the question of monetary non-neutrality (e.g., Amiti, Itskhoki, and Konings 2019; Sangani 2026).

[8] A distinct literature estimates the slope of the Phillips curve from cross-sectional regressions across regions (Hazell et al. 2022; Cerrato and Gitti 2022; Fitzgerald et al. 2024). Our statistic identifies the marginal-cost Phillips curve while this approach typically measures the output or unemployment based Phillips curve, which jointly identifies the slope of the marginal-cost Phillips curve and the response of marginal cost to the output gap.

2. Environment and Generalized Phillips Curve

This section lays out the environment common to the three models we consider and defines the object we want to recover, the generalized Phillips curve. Throughout, lower-case letters denote log deviations from a zero-inflation steady state.

2.1. Environment

Firms and marginal costs. A continuum of firms $i \in [0,1]$ produce differentiated varieties, aggregated with a CES aggregator with elasticity $\varepsilon > 1$, so that firm i faces the demand $Y_{it} = (P_{it}/P_t)^{-\varepsilon} Y_t$. Each firm produces with the technology $Y_{it} = A_{it} N_{it}^{\gamma}$, $\gamma \in (0,1]$, where N_{it} is labor hired at the economy-wide nominal wage W_t and A_{it} is idiosyncratic productivity. In log deviations, the nominal marginal cost of firm i is

$$mc_{it} = mc_t - \frac{a_{it}}{\gamma} + \frac{1-\gamma}{\gamma}\,(y_{it} - Y_t), \qquad mc_t \equiv w_t + \frac{1-\gamma}{\gamma}\,Y_t, \tag{1}$$

where mc_t is what we call *aggregate nominal marginal cost*. Real per-period profit at the nominal price P_{it} is:

$$\Pi_t(P_{it}) = \Big(\frac{P_{it}}{P_t}\Big)^{1-\varepsilon} Y_t - \frac{W_t}{P_t}\left[\frac{Y_t}{A_{it}}\Big(\frac{P_{it}}{P_t}\Big)^{-\varepsilon}\right]^{1/\gamma}. \tag{2}$$

The flexible-price—or frictionless desired price—maximizes (2) and is equal to a constant markup $\frac{\varepsilon}{\varepsilon-1}$ over the firm's own marginal cost. Solving for this quantity, we obtain

$$p_{it}^{*} = \omega\Big(mc_t - \frac{a_{it}}{\gamma}\Big) + (1-\omega)\,P_t, \qquad \omega \equiv \frac{1}{1+\varepsilon(1-\gamma)/\gamma} \in (0,1], \tag{3}$$

once the constant $\omega \ln \frac{\varepsilon}{\varepsilon-1}$ is dropped. Throughout, "desired price" is a shorthand for (3). With constant returns, $\omega = 1$, and the desired price is nominal marginal cost itself. With decreasing returns a firm that raises its price sells less and so produces at lower marginal cost, which damps the desired response to a cost shock by a factor $\omega < 1$. The coefficient ω acts as an index of *micro-level real rigidity*.[9]

[9] Every result below uses only the reduced form (3) together with a loss that is quadratic in the gap. Two other sources of micro-level real rigidities deliver the same reduced form: a firm-specific labor market whose wage the firm internalizes, $w_{it} - w_t = \varphi(n_{it} - N_t)$ with φ the inverse Frisch elasticity, which is isomorphic to (1) with $\frac{1-\gamma}{\gamma} = \varphi$; and Kimball demand, for which (3) holds to first order with ω set by the superelasticity.

Loss function. Let $x_{it} \equiv p_{it} - p_{it}^* = \ln P_{it} - \ln P_{it}^*$ denote the *price gap*. Expanding (2) to second order around the desired price, the per-period cost of a mispriced good is

$$\Pi_t(P_{it}^*) - \Pi_t(P_{it}) = \frac{1}{2}\, C_{it}\, x_{it}^2 + O(|x_{it}|^3), \qquad C_{it} = \frac{\varepsilon - 1}{\omega}\, Y_t \Big(\frac{P_{it}^*}{P_t}\Big)^{1-\varepsilon} \tag{4}$$

Dividing by C_{it} then measures every loss in units of the curvature: the per-period loss becomes $\frac{1}{2}x_{it}^2$, and the cost of adjusting, which depends on the adjustment technology introduced below, is measured in the same unit. Where that cost is a menu cost we take it to be proportional to frictionless revenue, ξC_{it}, so that in curvature units it is a constant ξ common across firms and dates. Using standard second-order arguments, equation (4) leads to the following objective function:

$$\mathbb{E}_0 \sum_{t \geq 0} \beta^t \Big[\frac{1}{2} x_{it}^2 + (\text{cost of adjusting}_{it})\Big], \tag{5}$$

Adjustment technologies. We consider two distinct adjustment technologies.

Calvo. This is model (C). With probability $1 - \alpha$, i.i.d. across firms and time, the firm may reset its price at no cost; otherwise its price is frozen.

Menu cost with free adjustments. The Gertler-Leahy model (GL) and canonical menu cost model (MC) both use this technology. The firm pays a menu cost $\xi_{it} \in \{0, \xi\}$ to change its price, with a free adjustment ($\xi_{it} = 0$) materializing with probability λ, i.i.d. across firms and time: the CalvoPlus specification of Nakamura and Steinsson (2010),[10] which nests Golosov and Lucas (2007) ($\lambda = 0$) and Calvo ($\xi \to \infty$). Whenever the firm reconsiders its price, the optimal policy has an Ss form: it adjusts if and only if its gap lies outside an inaction band $[-\bar{x}, \bar{x}]$ around the reset price (Appendix A.2.2 and B.4.1).

Idiosyncratic shocks. We decompose idiosyncratic productivity into two components,

$$-\frac{a_{it}}{\gamma} = z_{it} + \vartheta_{it}, \tag{6}$$

In this latter case, the curvature of the loss then depends on the firm's relative cost, so the reduction to the gap below holds only to second order. We provide numerical results for this case in Section 5.1.

[10] In Nakamura and Steinsson's specification the menu cost takes a low or a high value each period; we set the low value to zero, so that $\xi_{it} \in \{0, \xi\}$. We refer to this formulation as the canonical menu cost model (MC); Golosov and Lucas (2007) is the particular case where $\lambda = 0$.

so that

$$p_{it}^{*} = \omega(mc_t + z_{it} + \vartheta_{it}) + (1-\omega)\, P_t. \tag{7}$$

The component z_{it} is a generic productivity process. The component ϑ_{it} is a cost shifter that the econometrician observes and uses as a regressor; for instance in Herreño, Pinardon-Touati, and Thie (2026) it is a firm-specific input-cost shock. The split (6) is a statement about what the econometrician observes, not about how the firm behaves. Both components are independent driftless random walks:

$$\vartheta_{it} = \vartheta_{i,t-1} + \epsilon_{it}, \qquad \mathbb{E}[\epsilon_{it}] = 0, \quad \text{Var}(\epsilon_{it}) = \sigma_{\vartheta}^2 \tag{8}$$

$$z_{it} = z_{i,t-1} + \epsilon_{it}^{z}, \qquad \mathbb{E}[\epsilon_{it}^{z}] = 0, \quad \text{Var}(\epsilon_{it}^{z}) = \sigma_{z}^2. \tag{9}$$

where ϵ_{it} and ϵ_{it}^{z} are mutually independent, i.i.d. across firms and time, and independent of all adjustment draws at all dates; in particular, both are independent of the firm's predetermined state (its price, its gap and its adjustment history). The three models differ in the distributional assumptions on these shocks.

Calvo (*C*). ϵ_{it} and ϵ_{it}^{z} can have any distribution with the moments above.

Menu Cost (*MC*). ϵ_{it}^{z} has a symmetric, single-peaked and continuously differentiable density. For our baseline identification results, ϵ_{it} is Gaussian, and we discuss the non-Gaussian case in Appendix B.5.4.

Gertler-Leahy (*GL*). The driving shock is a large *turbulence* shock: with probability $1-\eta$, i.i.d. across firms and time, $z_{it} = z_{i,t-1} + \epsilon_{it}^{z}$ with $\omega\epsilon_{it}^{z} \sim U[-\varsigma/2, \varsigma/2]$. By contrast, innovations to mc_t, P_t and ϑ_{it} are small: they move the desired price by at most a first-order constant m over the horizon a price is expected to last, and $\varsigma > 4\bar{x} + 2m$, so that the band sits strictly inside the support of the turbulence shock. As explained below, this infrequent, uniform and large turbulence draw is what makes the model tractable.

With these assumptions, the desired price is a random walk in every model.[11] Then,

[11] The random-walk assumption is a simplification. In (C) and (GL) it can be relaxed for both z and ϑ and every analytical result goes through with modified formulas. In (MC), where the firm reconsiders its price every period, a non-random-walk shock makes the firm expect its desired price to mean-revert, so the reset gap and the band depend on the level of the shock and the gap is no longer the only state; for the identifying shock this is innocuous when its variance is small relative to the driving shock and the band. Appendix B.6

the firm's problem depends on $(mc_t, z_{it}, \vartheta_{it}, P_t)$ only through p^*_{it}. In the stationary cross-section the reset price is then the current desired price and the reset gap is zero.[12]

Aggregation. To first order the log price index is $P_t = \int_0^1 p_{it}\, di$, and since $\int (z_{it} + \vartheta_{it})\, di = 0$, $\int p^*_{it}\, di = \omega\, mc_t + (1-\omega) P_t$, so that $P_t = mc_t + \omega^{-1} \int x_{it}\, di$. Inflation is $\pi_t = P_t - P_{t-1}$. Real marginal cost is $mc^r_t \equiv mc_t - P_t$. The rest of the economy (labor supply, the Euler equation, monetary policy) determines mc^r_t and is only needed for the general-equilibrium exercises.

2.2. Pass-Through Jacobians and the Generalized Phillips Curve

We use sequence-space notation: bold symbols denote sequences starting at date 0, column vectors indexed by $t = 0, 1, \ldots$ (truncated at a horizon T in computations), $\mathbf{P} = (P_0, P_1, \ldots)'$, $\mathbf{mc} = (mc_0, mc_1, \ldots)'$, and so on. L and F are the lag and lead operators, square matrices of the same dimension with ones on the first subdiagonal and superdiagonal respectively, so that $(L\mathbf{x})_t = x_{t-1}$, $(L\mathbf{x})_0 = 0$ and $(F\mathbf{x})_t = x_{t+1}$; I is the identity matrix and $\mathbf{1} = (1, 1, \ldots)'$. For a Jacobian matrix J we write $J[t, s]$ for its entry in row t and column s, both indexed from 0, so that $[J\mathbf{1}]_h = \sum_s J[h, s]$ is the date-h response to a permanent unit shock.

Starting from the steady state, each model $m \in \{\mathrm{C}, \mathrm{GL}, \mathrm{MC}\}$ maps a perfect-foresight sequence of desired prices into a sequence of price levels through the aggregation of individual price-setting responses. Its first-order representation is the *desired-price pass-through Jacobian*

$$J^m_{P,p^*}[t, s] \equiv \frac{\partial P_t}{\partial p^*_s}, \qquad \mathbf{P} = J^m_{P,p^*}\, \mathbf{p}^*, \tag{10}$$

the response of the price level at t to a common shift of every firm's desired price at date s. This is a partial-equilibrium object: it treats the path of desired prices as exogenous and includes no feedback from prices to desired prices. By (7), holding the price level constant, a unit change in mc_s shifts every desired price by ω at date s, so the *nominal marginal cost*

details the three cases. Section 5.3 shows numerical results for a model with mean-reverting idiosyncratic process

[12] See Section 3.3.1 for (C), Appendix A.2 for (GL), Appendix B.4.1 for (MC).

pass-through Jacobian is

$$J^m_{P,mc}[t,s] \equiv \frac{\partial P_t}{\partial mc_s}\Big|_{\mathbf{P}} = \omega\, J^m_{P,p^*}[t,s]. \tag{11}$$

Using $\mathbf{p}^* = \omega\,\mathbf{mc} + (1-\omega)\mathbf{P}$ and $\mathbf{mc}^r = \mathbf{mc} - \mathbf{P}$, the price level solves the fixed point $\mathbf{P} = J^m_{P,p^*}(\omega\,\mathbf{mc} + (1-\omega)\mathbf{P}) = J^m_{P,p^*}(\omega\,\mathbf{mc}^r + \mathbf{P})$, so the *real marginal cost pass-through Jacobian* writes:

$$J^m_{P,mc^r} = (I - J^m_{P,p^*})^{-1}\omega\, J^m_{P,p^*} = \omega\,(I - J^m_{P,p^*})^{-1} J^m_{P,p^*}, \tag{12}$$

Finally, we define the *generalized Phillips curve* Jacobian, as:

$$J^m_{\pi,mc^r} = (I - L)\, J^m_{P,mc^r}, \tag{13}$$

Given J^m_{π,mc^r}, inflation is pinned down by the path of real marginal cost, $\boldsymbol{\pi} = J^m_{\pi,mc^r}\mathbf{mc}^r$, and hence by any demand-side block that determines $\mathbf{mc}^r$. Equations (12)–(13) are pure linear algebra given J^m_{P,p^*} and ω, with J^m_{P,p^*} the only model-specific object.

3. A Robust Statistic for the Generalized Phillips Curve

3.1. The Statistic

Consider the following regression in the stationary cross-section of firms, for horizons $h = 0, 1, 2, \ldots$ (the statistic uses $h = 0$ and $h = 1$):

$$p_{it+h} - p_{i,t-1} = a_h + b_h \epsilon_{it} + u_{it} \tag{14}$$

b_h is the population local projection of cumulative price changes on the idiosyncratic innovation. Then, define:

$$\phi^{\text{IRF}} \equiv \frac{b_1}{b_0} - 1 \tag{15}$$

$$\kappa^{\text{IRF}} \equiv \frac{b_0(1 - \beta\phi^{\text{IRF}})}{\phi^{\text{IRF}}} \tag{16}$$

Let $\Theta_\beta \equiv (I - \beta F)^{-1}$. Our estimate of the generalized Phillips curve is then:

$$J^{\text{IRF}}_{\pi,mc^r} = \kappa^{\text{IRF}}\Theta_\beta = \kappa^{\text{IRF}} \begin{bmatrix} 1 & \beta & \beta^2 & \beta^3 & \cdots \\ 0 & 1 & \beta & \beta^2 & \cdots \\ 0 & 0 & 1 & \beta & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{17}$$

There are two key steps for understanding the identification of the generalized Phillips curve. First, we show that the micro IRF b_h is equal to the response of the aggregate price level to a permanent unit shock to nominal marginal cost, $[J^m_{P,mc}\mathbf{1}]_h$ at every horizon h. Second, we show that the first two entries of $J^m_{P,mc}\mathbf{1}$—identified by b_0 and b_1—allow to provide a highly accurate approximation of the generalized Phillips curve J^m_{π,mc^r} in a wide class of models.

3.2. From the Cross-Sectional to the Aggregate Pass-Through

In every model m, the population local projection of cumulative price changes on the idiosyncratic innovation is given by:

$$b^m_h \equiv \frac{\text{Cov}\left(p_{i,t+h} - p_{i,t-1},\ \epsilon_{it}\right)}{\text{Var}(\epsilon_{it})}, \qquad h = 0, 1, 2, \ldots, \tag{18}$$

We write $b^m = (b^m_0, b^m_1, \ldots)'$ and call it the *micro IRF*.

Proposition 1 (Micro–macro bridge). *In each model, under the conditions listed below, the micro IRF equals the response of the aggregate price level to a permanent unit shock to nominal marginal cost:*

$$b^m_h = [J^m_{P,mc}\mathbf{1}]_h \qquad \textit{for all } h \geq 0. \tag{19}$$

The conditions, beyond the independence assumptions on the innovations maintained after (9)*, are: none in* (C)*; the small-shock assumptions of Section 2.1 in* (GL)*; and, in* (MC)*,* $\epsilon_{it} \sim N(0, \sigma^2_\vartheta)$.

Proof (*main argument, full proof in Appendix B.5*). Fix t and a firm. Let S be its predetermined state, ϵ_{it} its identifying innovation, and $\mathcal{D}$ the other random draws that affect its

prices through $t+h$; the cumulative price change is then a deterministic function $f_h(\epsilon_{it}; S, \mathcal{D})$ of the innovation. The proof has two steps.

The first step shows that the aggregate response is the cross-sectional average derivative of the price path with respect to the firm's *own* shifter,

$$[J^m_{P,mc}\mathbf{1}]_h = \frac{d}{d\delta}\,\mathbb{E}_{S,\epsilon,\mathcal{D}}[f_h(\epsilon_{it}+\delta; S, \mathcal{D})]\Big|_{\delta=0}, \tag{20}$$

in every model. Under the random walk, the shift δ raises the desired price by $\omega\delta$ at every $s \geq 0$ and changes nothing else. Since the firm's problem depends on $(mc_t, z_{it}, \vartheta_{it}, P_t)$ only through p^*_{it}, and the price level is held fixed in (11), this is, from the firm's point of view, an unanticipated permanent increase in mc by δ. Averaging over the stationary cross-section gives the response of the aggregate price level to a permanent unit nominal-cost shock with the price level fixed.[13]

The second step shows that the regression coefficient b^m_h equals the right-hand side of (20). The regression weights each firm's price change by its own innovation; the average derivative weights the local responses of f_h by the population density of the innovation. The two coincide: (i) for every response function f_h when the innovation is Gaussian, by Stein's identity, or (ii) for any distribution when the response (averaged over the other draws $\mathcal{D}$) is affine in the innovation, which is the case in (C) and (GL). □

Corollary 1 (Pass-through and the surviving gap). *In every model,*

$$b^m_h = \omega + \frac{\text{Cov}\,(x_{i,t+h}, \epsilon_{it})}{\sigma^2_\vartheta}. \tag{21}$$

Proof. Write $p_{i,t+h} - p_{i,t-1} = (p^*_{i,t+h} - p^*_{i,t-1}) + x_{i,t+h} - x_{i,t-1}$. With ϑ a random walk and ϵ_{it} independent of all other innovations and of the aggregates, $\text{Cov}(p^*_{i,t+h} - p^*_{i,t-1}, \epsilon_{it}) = \omega\sigma^2_\vartheta$ by (7), and $x_{i,t-1}$ is predetermined so $\text{Cov}(x_{i,t-1}, \epsilon_{it}) = 0$. □

Pass-through at horizon h falls short of ω, the pass-through of the shock into the desired price, exactly to the extent that the price gap h periods after the shock still carries the shock. This identity holds in every model and is the device we use to explain the shape of the micro IRF. It also shows what the long-run pass-through measures: with a random-walk

[13] Caballero and Engel (2007) offer a related but different result: they relate the macro IRF to a common shock to the average of firms' marginal responses to that same shock; we relate it to the average of firms' marginal responses to their own micro shock.

shifter, $b_h^m \to \omega$, the desired-price pass-through coefficient, irrespective of the adjustment technology.

Extensions. We discuss the case where the identifying shock is not Gaussian in the (MC) model in Appendix B.5.4 and the case of non-random walk shocks in Appendix B.6.

What is not identified. The micro IRF identifies the vector $J^m_{P,mc}\mathbf{1}$—the response to an unanticipated permanent cost shock with the price level held fixed—and not the matrix $J^m_{P,mc}$: idiosyncratic shocks are never anticipated, so the columns describing responses to anticipated future cost changes are outside the regression's reach. In addition, without further structure, it does not separately identify ω from J^m_{P,p^*} at any finite horizon. Since the generalized Phillips curve (13) depends on the entire matrix J^m_{P,p^*}, the micro IRF does not identify it in general.

3.3. From the Pass-Through Matrix to the Generalized Phillips Curve

This section shows how only the first two entries of $J^m_{P,mc}\mathbf{1}$ allow to construct a highly robust estimate of the generalized Phillips curve across price setting models. To build intuition, we start from the Calvo model, where the two entries of $J^m_{P,mc}\mathbf{1}$ are sufficient to recover the entire generalized Phillips curve exactly. We then show that the same two entries are exactly sufficient in the (GL) model, which allows to provide intuition for our key result that—even if not exact—our statistic provides an extremely robust estimate of the generalized Phillips curve in the general class of (MC) models.

3.3.1. Calvo and the IRF Statistics

For any non-adjustment probability $\alpha \in (0,1)$ define

$$\Upsilon(\alpha) \equiv (I-\alpha L)^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ \alpha & 1 & 0 & \cdots \\ \alpha^2 & \alpha & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \qquad \Theta(\alpha) \equiv (I-\beta\alpha F)^{-1} = \begin{bmatrix} 1 & \beta\alpha & (\beta\alpha)^2 & \cdots \\ 0 & 1 & \beta\alpha & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

A firm that resets at t keeps its price until its next adjustment opportunity, so it minimizes $\mathbb{E}_t \sum_{s\geq 0}(\beta\alpha)^s \frac{1}{2}(p_{it}^{\text{new}} - p_{i,t+s}^*)^2$ and sets

$$p_{it}^{\text{new}} = (1-\beta\alpha)\,\mathbb{E}_t \sum_{s\geq 0}(\beta\alpha)^s\, p_{i,t+s}^*, \qquad \text{or} \qquad \mathbf{p}_i^{\text{new}} = (1-\beta\alpha)\,\Theta(\alpha)\,\mathbf{p}_i^*, \tag{22}$$

where $\mathbf{p}_i^* = \omega(\mathbf{mc} + \mathbf{z}_i + \boldsymbol{\vartheta}_i) + (1-\omega)\mathbf{P}$ by (7) and, for the idiosyncratic components, the sequence is the expected path. $\Theta(\alpha)$ maps the sequence of desired prices into the sequence of reset prices. The expected price of the firm weights reset prices across vintages by the survival probabilities,

$$\bar{p}_{it} = (1-\alpha)\sum_{j\geq 0}\alpha^j\, p_{i,t-j}^{\text{new}}, \qquad \text{or} \qquad \bar{\mathbf{p}}_i = (1-\alpha)\,\Upsilon(\alpha)\,\mathbf{p}_i^{\text{new}}. \tag{23}$$

$\Upsilon(\alpha)$ maps reset prices into the price level by aggregating over surviving vintages. Then,

$$\bar{\mathbf{p}}_i = J_{P,p^*}^{\text{C}}(\alpha)\,\mathbf{p}_i^*, \qquad J_{P,p^*}^{\text{C}}(\alpha) \equiv (1-\alpha)(1-\beta\alpha)\,\Upsilon(\alpha)\Theta(\alpha). \tag{24}$$

Aggregating across firms, the idiosyncratic components wash out and $\mathbf{P} = J_{P,p^*}^{\text{C}}(\alpha)(\omega\mathbf{mc} + (1-\omega)\mathbf{P})$: the desired-price pass-through of the Calvo model is $J_{P,p^*}^{\text{C}}(\alpha)$, and its nominal marginal cost pass-through Jacobian is $J_{P,mc}^{\text{C}}(\alpha) = \omega\, J_{P,p^*}^{\text{C}}(\alpha)$. Two properties of $J_{P,p^*}^{\text{C}}(\alpha)$ drive everything that follows.

LEMMA 1. *Let $\kappa(\alpha) \equiv (1-\alpha)(1-\beta\alpha)/\alpha$. Then*

(i) $[J_{P,p^*}^{\text{C}}(\alpha)\mathbf{1}]_h = 1 - \alpha^{h+1}$;

(ii) $(I-L)(I - J_{P,p^*}^{\text{C}}(\alpha))^{-1} J_{P,p^*}^{\text{C}}(\alpha) = \kappa(\alpha)\Theta_\beta$, *so that by* (12)–(13) $J_{\pi,mc^r}^{\text{C}}(\alpha) = \omega\,\kappa(\alpha)\Theta_\beta$.

Proof: Appendix B.1.1.

Property (ii) is the New Keynesian Phillips curve $\pi_t = \omega\kappa(\alpha)\, mc_t^r + \beta\pi_{t+1}$ written in sequence space $\boldsymbol{\pi} = \omega\,\kappa(\alpha)\,\Theta_\beta\,\mathbf{mc}^r$.

PROPOSITION 2 (The IRF statistics in the Calvo model). *Suppose that in any model m the desired-price pass-through Jacobian has the Calvo structure, $J_{P,p^*}^m = J_{P,p^*}^{\text{C}}(\alpha)$ for some $\alpha \in (0,1)$, with desired-price pass-through $\omega \in (0,1]$, and that Proposition 1 applies. Then*

(*i*) $b_h^m = \omega(1-\alpha^{h+1})$, *so that*

$$\phi^{\text{IRF},m} \equiv \frac{b_1^m}{b_0^m} - 1 = \alpha; \tag{25}$$

(*ii*) *the generalized Phillips curve is* $J^m_{\pi,mc^r} = \kappa^{\text{IRF},m}\,\Theta_\beta$, *i.e.* $\pi_t = \kappa^{\text{IRF},m}\,mc_t^r + \beta\pi_{t+1}$, *with*

$$\kappa^{\text{IRF},m} \equiv \frac{b_0^m\,(1-\beta\phi^{\text{IRF},m})}{\phi^{\text{IRF},m}}. \tag{26}$$

Hence, given β, *the first two entries of the micro IRF determine the entire generalized Phillips curve.*

Proof. (i) follows from Proposition 1 and Lemma 1(i). (ii) follows from Lemma 1(ii), $J^m_{\pi,mc^r} = \omega\kappa(\alpha)\Theta_\beta$, with $\omega\kappa(\alpha) = \omega(1-\alpha)(1-\beta\alpha)/\alpha = b_0(1-\beta\alpha)/\alpha$ by (i). □

This works because of two things. b_0 captures the first entry of the passthrough Jacobian. The mapping from the on-impact passthrough to the Phillips curve only depends on α (given β), which is identified by the ratio b_1/b_0. The two entries of the micro IRF are therefore sufficient to recover the entire generalized Phillips curve.

The micro-to-macro mapping is a fixed-point problem: firms reset their prices taking the aggregate price level as given. The strength of the feedback from prices to desired prices is governed by $1-\alpha$, the degree of price flexibility (and does not depend on ω, since the price level enters the desired price with total weight 1). The micro IRF identifies α because, by Corollary 1, $\omega - b_h = -\text{Cov}(x_{i,t+h}, \epsilon_{it})/\sigma^2_\vartheta = \omega\alpha^{h+1}$: the shortfall of pass-through from its long-run value ω at horizon h is ω times the probability that the shock still sits in the firm's price gap. This shortfall decays at the geometric rate α, and the ratio $b_1/b_0 = 1+\alpha$ reads it off because real rigidity scales every entry of the micro IRF by the same ω. If prices are very sticky, many firms cannot adjust in the first period, so a larger remaining mass of firms still wants to adjust between $t=0$ and $t=1$. The level b_0 then identifies ω.[14]

Remark 1 (Persistent shifters). *With* $\vartheta_{it} = \rho\vartheta_{i,t-1} + \epsilon_{it}$ *and* $J^m_{P,p^*} = J^{\text{C}}_{P,p^*}(\alpha)$, *we instead have:*

$$b_0 = \omega\,\frac{(1-\alpha)(1-\beta\alpha)}{1-\beta\alpha\rho}, \qquad \frac{b_1}{b_0} = \alpha + \rho,$$

[14] Given a third entry the system is overidentified, $b_2 = \omega(1-\alpha^3)$, which provides a specification test of the Calvo structure. Note that the micro IRF does not identify β: under a random walk no entry depends on β, and under an AR(1) shifter β enters only through the level factor $(1-\beta\alpha)/(1-\beta\alpha\rho)$ of Remark 1, which is confounded with ω.

so that $\alpha = b_1/b_0 - \rho$ *and* $\kappa = \omega\kappa(\alpha) = b_0(1-\beta\alpha\rho)/\alpha$*, with* ρ *estimated from the autocorrelation of the shifter. The random-walk formulas* (25)–(26) *are the case* $\rho = 1$*.*

We can already note that the Calvo structure of $J^{\mathrm{C}}_{P,p^*}(\alpha)$ is not tied to the Calvo lottery. In the Rotemberg model, where every price changes every period at a convex cost, the desired-price pass-through is exactly $J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})$ with α_{R} the stable root of the firm's Euler equation, so Proposition 2 applies verbatim and the statistic is exact while the frequency of price changes is one (Appendix B.2).

3.3.2. Gertler–Leahy

In the GL economy the decision to adjust is state dependent, so there is no structural non-adjustment probability that a Calvo econometrician could target. Nevertheless, the desired-price pass-through Jacobian is exactly Calvo, with the probability $\eta_\lambda \equiv (1-\lambda)\eta$ that a price is neither hit by turbulence nor freely reconsidered in place of α, and our statistic recovers η_λ (set-up in Appendix A.2, proofs in Appendix B.3).

Reset behavior. A key result is that a firm reconsiders its price only when hit by turbulence or when it draws a free adjustment.[15] A price therefore lasts until the next hit or free draw, an event of per-period probability $1-\eta_\lambda$. After a hit the firm adjusts if and only if its would-be gap leaves the band $[-\bar{x},\bar{x}]$, $\bar{x} = \sqrt{2(1-\eta_\lambda\beta)\xi}$; after a free draw it always adjusts.

A firm that adjusts at t faces the same trade-off as a Calvo resetter, since its price survives each period with probability η_λ regardless of its gap: its reset price is the Calvo formula (22) with η_λ in place of α, evaluated along the path with no further hit,

$$p_{it}^{\text{new}} = \omega z_{it} + (1-\eta_\lambda\beta)\,\mathbb{E}_t \sum_{s\geq 0} (\eta_\lambda\beta)^s \Big[\omega(mc_{t+s} + \vartheta_{i,t+s}) + (1-\omega)P_{t+s}\Big], \tag{27}$$

where z_{it} enters separately because it is frozen until the next hit. Let $\bar{p}_{it}$ and $\bar{p}_{it}^{\text{new}}$ denote the expectations of p_{it} and p_{it}^{new} over the turbulence process (hits and draws), given the

[15] This follows from the decision-cost assumption (A4) of Appendix A.2: a firm that is neither hit nor free does not pay the cost and keeps its price. The small-shock assumptions (A1)–(A3) make (A4) consistent, since the gain from reconsidering without a hit is second order while the gain after a hit is not, and keep the band strictly inside the support of the turbulence draw.

paths of mc, P and ϑ_i. Turbulence draws being mean zero,

$$\bar{p}_{it}^{\text{new}} = (1-\eta_\lambda\beta)\,\mathbb{E}_t \sum_{s\geq 0}(\eta_\lambda\beta)^s \Big[\omega(mc_{t+s}+\vartheta_{i,t+s})+(1-\omega)P_{t+s}\Big]. \tag{28}$$

Aggregation. Consider a firm hit at t without a free draw. Its would-be gap, $y_{it} \equiv p_{i,t-1} - p_{it}^{\text{new}}$, is the gap after its last hit, inside the band, plus the drift of the reset price since then, bounded by m under (A2), minus the turbulence draw $\omega\epsilon_{it}^z$, uniform with support width $\varsigma > 4\bar{x} + 2m$: y_{it} is therefore uniform on an interval of width ς containing the band. The firm keeps its price if $|y_{it}| \leq \bar{x}$, probability $2\bar{x}/\varsigma$, and resets otherwise, so

$$\begin{aligned}\mathbb{E}[p_{it} \mid \text{hit}] &= \Pr(\text{adjust} \mid \text{hit})\,\mathbb{E}[p_{it} \mid \text{hit, adjust}] + \Pr(\text{keep} \mid \text{hit})\,\mathbb{E}[p_{it} \mid \text{hit, keep}] \\ &= \Big(1-\frac{2\bar{x}}{\varsigma}\Big)p_{it}^{\text{new}} + \frac{2\bar{x}}{\varsigma}\Big(p_{it}^{\text{new}} + \mathbb{E}[y_{it} \,\big|\, |y_{it}| \leq \bar{x}]\Big) = p_{it}^{\text{new}},\end{aligned} \tag{29}$$

since $\mathbb{E}[y_{it} \mid |y_{it}| \leq \bar{x}] = 0$: non-adjusters are spread symmetrically around the reset price, so the expected price of a hit firm equals its reset price whether or not it adjusts (Lemma B.2). Free adjusters are at the reset price by construction. Hence

$$\begin{aligned}\bar{p}_{it} &= \underbrace{\lambda\,\bar{p}_{it}^{\text{new}}}_{\text{free adjusters}} + \underbrace{(1-\lambda)(1-\eta)\,\bar{p}_{it}^{\text{new}}}_{\text{hit, no free draw}} + \underbrace{\eta_\lambda\,\bar{p}_{i,t-1}}_{\text{neither}} \\ &= \eta_\lambda\,\bar{p}_{i,t-1} + (1-\eta_\lambda)\,\bar{p}_{it}^{\text{new}},\end{aligned} \tag{30}$$

the Calvo law (23) with hazard $1-\eta_\lambda$ and reset price (28): adjustment is state dependent firm by firm, but in expectation every price behaves as if reset with probability $1-\eta_\lambda$ regardless of its gap. Integrating across firms, $\mathbf{P} = J^{\text{C}}_{P,p^*}(\eta_\lambda)[\omega\mathbf{mc} + (1-\omega)\mathbf{P}]$.

Proposition 3 (Gertler–Leahy). *In the GL economy,*

(*i*) *The desired-price pass through Jacobian is Calvo:* $J^{\text{GL}}_{P,p^*} = J^{\text{C}}_{P,p^*}(\eta_\lambda)$*, and hence* $J^{\text{GL}}_{\pi,mc^r} = J^{\text{C}}_{\pi,mc^r}(\eta_\lambda)$*;*

(*ii*) *The IRF statistics are exact:* $b_h^{\text{GL}} = \omega(1-\eta_\lambda^{h+1})$*;* $\phi^{\text{IRF,GL}} = \eta_\lambda$*; and* $J^{\text{GL}}_{\pi,mc^r} = \kappa^{\text{IRF,GL}}\Theta_\beta$*;*

(*iii*) *The synthetic hazard* $1-\phi^{\text{IRF,GL}}$ *exceeds the actual frequency of adjustment:* $\text{freq} = \lambda + (1-\lambda)(1-\eta)(1-2\bar{x}/\varsigma) < 1-\eta_\lambda$.

Proof: Appendix B.3.2.[16]

[16] The proof relies on Proposition 1 applying. Note that this is not the case if, instead of the shifter inno-

Despite the model not being Calvo, the statistics b_0 and $\phi^{\text{IRF,GL}}$ recover the correct structural objects, even if they depend on η_λ as opposed to the actual frequency of adjustment freq. First, $b_0 = \omega(1-\eta_\lambda)$. The shifter innovation moves the desired price by $\omega\epsilon_{it}$ but never triggers a reconsideration; the price moves only on a hit or a free draw, with probability $1-\eta_\lambda$, and in that event the expected price is the reset price whether or not the firm adjusts, by (29), which moves one for one with $\omega\epsilon_{it}$. The expected price change at impact is thus $(1-\eta_\lambda)\omega\epsilon_{it}$. Here, selection means that firms that don't reset had the right price on average, and so the loading of the actual price on the desired price moved by the shock is higher than the actual frequency of adjustment. Second, the statistic $\phi^{\text{IRF,GL}}$ recovers η_λ and not freq. By Corollary 1, $\omega - b_h = -\operatorname{Cov}(x_{i,t+h}, \epsilon_{it})/\sigma^2_\vartheta$: the shortfall of pass-through from ω is the part of the shock still sitting in the firm's gap, in expectation over hits and draws. After a hit or a free draw the shock is fully absorbed in expectation, whether or not the firm adjusts; so it survives in the expected gap at horizon h only if neither event has occurred in $h+1$ periods, probability η_λ^{h+1}. Thus $\omega - b_h = \omega\eta_\lambda^{h+1}$, and $\phi^{\text{IRF,GL}} = \eta_\lambda$.

In the (GL) model, the assumptions ensure that the whole Jacobian is exactly Calvo.[17] However, the insights from this discussion carry over to the general (MC) model: our statistic approximates the true Jacobian even when the economy that generates it is not Calvo, because it measures the survival of the shock in the *expected* price, not the adjustment hazard of realized prices.

3.3.3. Canonical Menu Cost

In the canonical menu cost model, J^{MC}_{P,p^*} does not have the Calvo structure. We proceed in three steps: what b_0 identifies; what the term structure of the micro IRF identifies *exactly*, namely the average virtual survival function of Auclert et al. (2024)'s two-margin representation, of which $\phi^{\text{IRF,MC}}$ is the one-period-ahead value; and why extrapolating this object geometrically, which is all the Calvo structure adds, is accurate.

---

vation ϵ_{it}, the econometrician regresses on the turbulence innovation ϵ^z_{it} itself. See discussion in Appendix B.5.6.

[17] That is because the backward and forward dynamics are: the law of expected prices (30) has hazard $1-\eta_\lambda$ because reconsideration events, and the expected price conditional on them, do not depend on the gap; the reset price (27) discounts at $\eta_\lambda\beta$ because the new price survives until the next hit or free draw, again independently of the gap.

Policy function. The firm reconsiders its price every period, and its pre-adjustment gap is $x_{i,t-1} - \omega(\epsilon^z_{it} + \epsilon_{it})$. In the steady state the reset gap is $x^* = 0$ and the band is symmetric (the menu cost counterpart of the martingale property: the desired-price innovation is a symmetric driftless random walk and nothing is anticipated); the pre-adjustment gap has a stationary density $g(x)$, and the frequency of price adjustment is $\text{freq} = 1 - (1 - \lambda) \int_{-\bar{x}}^{\bar{x}} g(x)\, dx$.

On-impact pass-through. By Proposition 1 and Appendix B.4.2,

$$b_0^{\text{MC}} = [J^{\text{MC}}_{P,mc}\mathbf{1}]_0 = \omega[\text{freq} + 2(1-\lambda)\, g(\bar{x})\, \bar{x}]. \tag{31}$$

The on-impact pass-through, measured by the regression coefficient b_0^{MC}, encodes the selection effect. Selection is the dependence of the adjustment decision on the gap: unless it gets a free adjustment, a firm adjusts if and only if its gap lies outside the band, so adjusters are the firms with the largest gaps, not a random sample of firms. b_0 contains it mechanically, because the shock displaces every gap and thereby pushes the firms sitting just inside the edges of the band into adjustment, and the covariance weights these recruited adjusters by their full price change, $\bar{x}$, rather than by the marginal displacement that recruited them. Therefore, the impact pass-through exceeds ω times the frequency (the Calvo-world coefficient).

Term structure of the pass-through matrix. The building block is a time-dependent (TD) model in which a price set at date t is still in place at $t+s$ with probability Φ_s ($\Phi_0 = 1$, Φ_s decreasing):

$$J^{\text{TD}}_{P,p^*}(\Phi) = \frac{1}{(\sum_{s\geq 0}\Phi_s)(\sum_{s\geq 0}\beta^s\Phi_s)}\, \Upsilon^{\Phi}\, \Theta^{\Phi}, \tag{32}$$

where Υ^{Φ} is lower-triangular Toeplitz with entries Φ_{t-s} and Θ^{Φ} is upper-triangular Toeplitz with entries $\beta^{s-t}\Phi_{s-t}$; the Calvo model is the case $\Phi_s = \alpha^s$, for which $J^{\text{TD}}_{P,p^*}(\Phi) = J^{\text{C}}_{P,p^*}(\alpha)$. Auclert et al. (2024) prove that the desired-price pass-through of the canonical menu cost model is a convex combination of two TD Jacobians,

$$J^{\text{MC}}_{P,p^*} = \varrho_e\, J^{\text{TD}}_{P,p^*}(\Phi^e) + \varrho_i\, J^{\text{TD}}_{P,p^*}(\Phi^i), \qquad \varrho_e + \varrho_i = 1, \qquad J^{\text{MC}}_{P,mc} = \omega\, J^{\text{MC}}_{P,p^*}, \tag{33}$$

whose *virtual survival functions* Φ^e and Φ^i encode the extensive margin (movements of the adjustment bands) and the intensive margin (movements of the reset price).[18]

LEMMA 2 (TD partial sums). *For any TD model,* $[J^{\text{TD}}_{P,p^*}(\Phi)\mathbf{1}]_h = \sum_{s=0}^{h} \Phi_s / \sum_{s\geq 0} \Phi_s$.

PROOF. Every row of Θ^Φ sums to $\sum_s \beta^s \Phi_s$, so $\Theta^\Phi \mathbf{1} = (\sum_s \beta^s \Phi_s)\mathbf{1}$; and $[\Upsilon^\Phi \mathbf{1}]_h = \sum_{s=0}^{h} \Phi_s$. □

The permanent-shock IRF of a TD model is the normalized partial sum of its survival function. Applying Lemma 2 to the mixture (33) and using Proposition 1, $b_h = \omega[J^{\text{MC}}_{P,p^*}\mathbf{1}]_h$,

$$b_h^{\text{MC}} = b_0^{\text{MC}} \sum_{s=0}^{h} \bar{\Phi}_s, \qquad \bar{\Phi}_s \equiv m_e\, \Phi_s^e + m_i\, \Phi_s^i \tag{34}$$

with weights $m_e = 2(1-\lambda)g(\bar{x})\bar{x}/(b_0/\omega)$ and $m_i = \text{freq}/(b_0/\omega)$ summing to one.

Therefore, the shape of the micro IRF identifies the average virtual survival function $\bar{\Phi}_h$ horizon by horizon:

$$\frac{b_h^{\text{MC}} - b_{h-1}^{\text{MC}}}{b_0^{\text{MC}}} = \bar{\Phi}_h, \quad h \geq 1. \tag{35}$$

In particular, our statistic recovers a structural object of the menu cost model exactly:

$$\phi^{\text{IRF,MC}} = \frac{b_1^{\text{MC}}}{b_0^{\text{MC}}} - 1 = \bar{\Phi}_1, \tag{36}$$

the average one-period survival of a marginal price gap. What the Calvo structure adds is only the geometric extrapolation of $\bar{\Phi}_1$ to later horizons, to which we turn below.

Like b_0^{MC}, $\phi^{\text{IRF,MC}}$ internalizes selection effects. It is the one-period survival of a marginal *gap*, not of a price spell: a firm that carries the shock past impact has a larger gap and is disproportionately likely to be pushed across a threshold next period, so the shock content of prices mean-reverts faster than price spells end, and $\phi^{\text{IRF,MC}}$ falls short of $1 - \text{freq}$ (Appendix B.4.2). The ratio reads the decay of the shock in the expected price—the object the Jacobian is built from—which selection accelerates; the frequency counts spells,

[18] Let $E^s(x)$ denote the steady-state expectation of the gap of a firm s periods ahead, given a gap x today ($E^0(x) = x$; the firm may adjust in between). Then $\Phi_s^e = E^s(\bar{x})/\bar{x}$, the surviving fraction of a marginal gap that starts at the adjustment threshold, and $\Phi_s^i = E^{s\prime}(0)$, the sensitivity of the expected future gap to the current gap at the reset point; the mixture weights are $\varrho_e = 2(1-\lambda)g(\bar{x})\bar{x}\sum_s \Phi_s^e$ and $\varrho_i = \text{freq}\sum_s \Phi_s^i$. See Auclert et al. (2024).

which selection leaves untouched. The gap between the two, $(1-\text{freq})-\phi^{\text{IRF,MC}}$, is therefore a model-free measure of the strength of the selection effect, a feature we document numerically below.

The Calvo extrapolation. Our IRF statistics assembles the generalized Phillips curve $J^{\text{IRF,MC}}_{\pi,mc^r}=J^{\text{C}}_{\pi,mc^r}(\phi^{\text{IRF,MC}})=\kappa^{\text{IRF,MC}}\Theta_\beta$ with first entry $\kappa^{\text{IRF,MC}}=b_0^{\text{MC}}(1-\beta\phi^{\text{IRF,MC}})/\phi^{\text{IRF,MC}}$. It takes the identified pair $([J^{\text{MC}}_{P,mc}\mathbf{1}]_0,\bar\Phi_1)$ and extrapolates the survival function geometrically, $\bar\Phi_s\approx(\phi^{\text{IRF,MC}})^s$. The extrapolation involves two approximations:

(i) *Along the permanent-shock IRF, horizons $h\geq 2$.* With the levels b_0^{MC} and b_1^{MC} matched, the error is

$$b_0^{\text{MC}}\sum_{s=2}^{h}(\bar\Phi_s-(\phi^{\text{IRF,MC}})^s),$$

the geometric extrapolation set against the true average survival function. It is small because, as Auclert et al. (2024) document numerically, the virtual hazards of the two margins converge to the same constant and deviate from it in *opposite* directions—the extensive margin from above, the intensive margin from below—so the deviations roughly offset in the average $\bar\Phi$, leaving it close to geometric. This approximation is also testable: $\bar\Phi_h=(\phi^{\text{IRF,MC}})^h$ at $h\geq 2$ is a geometric-decay restriction, and estimated deviations from it measure the extrapolation error directly.

(ii) *From the IRF to the matrix.* When each margin's survival function is close to the *same* geometric—which is what the offsetting virtual hazards deliver in practice—the shared $\Upsilon^\Phi\Theta^\Phi$ architecture of (32) transfers the proximity to every entry of the Jacobian.

A bound. For an infinite matrix A and a reporting horizon $H\geq 1$, let

$$\|A\|_H\equiv\max_{0\leq t\leq H}\sum_{s\geq 0}\big|A[t,s]\big| \tag{37}$$

denote the maximum absolute row sum over the first $H+1$ rows; it satisfies $|[A\mathbf{1}]_t|\leq\|A\|_H$ for $t\leq H$.

Assumption 1. *There exist $\alpha^\star\in(0,1)$ and $\bar e\geq 0$ with $\|J^{\text{MC}}_{P,p^*}-J^{\text{C}}_{P,p^*}(\alpha^\star)\|_H\leq\bar e$ and $\bar e<1-\alpha^\star$.*

The assumption says that *some* Calvo desired-price pass-through Jacobian is close to that of the menu cost model, the numerical finding of Auclert et al. (2024). Define the approx-

imation error $E \equiv J^{\text{MC}}_{P,p^*} - J^{\text{C}}_{P,p^*}(\alpha^\star)$. The proposition below then bounds the error of the Calvo Jacobian assembled from $\phi^{\text{IRF,MC}}$.

Proposition 4 (Menu cost). *Under the conditions of Proposition 1 and Assumption 1:*

(*i*) *the desired-price pass-through Jacobian satisfies*

$$\|J^{\text{MC}}_{P,p^*} - J^{\text{C}}_{P,p^*}(\phi^{\text{IRF,MC}})\|_H \le \bar{e} + M_H(I_\alpha)\,\frac{(2+\alpha^\star)\,\bar{e}}{1-\alpha^\star-\bar{e}} \;\equiv\; B_H, \tag{38}$$

where $I_\alpha \equiv [\min\{\alpha^\star, \phi^{\text{IRF,MC}}\}, \max\{\alpha^\star, \phi^{\text{IRF,MC}}\}]$ *and* $M_H(I_\alpha) \equiv \sup_{\alpha\in I_\alpha}\|\partial J^{\text{C}}_{P,p^*}(\alpha)/\partial\alpha\|_H$*, a finite number determined by* β*,* H *and* I_α *whenever* $I_\alpha \subset (0,1)$*; hence the nominal Jacobians satisfy* $\|J^{\text{MC}}_{P,mc} - \omega J^{\text{C}}_{P,p^*}(\phi^{\text{IRF,MC}})\|_H \le \omega B_H$ *with the true* ω*;*

(*ii*) *on the system truncated at the reporting horizon* H*, with* $\|\cdot\|$ *the maximum-absolute-row-sum norm (submultiplicative there) and the truncated* $I - J^{\text{MC}}_{P,p^*}$ *invertible, the generalized Phillips curves built from* (12)–(13) *with the true real rigidity* ω *satisfy*

$$\|J^{\text{MC}}_{\pi,mc^r} - J^{\text{C}}_{\pi,mc^r}(\phi^{\text{IRF,MC}})\| \le 2\,\omega\,\|(I - J^{\text{C}}_{P,p^*}(\phi^{\text{IRF,MC}}))^{-1}\|\,\|(I - J^{\text{MC}}_{P,p^*})^{-1}\|\,B_H; \tag{39}$$

Proof. See Appendix B.4.3. □

Part (ii) deserves a comment. Every entry of J^m_{π,mc^r} depends on the *whole* pass-through matrix: $(I - J^m_{P,p^*})^{-1}J^m_{P,p^*} = \sum_{k\ge1}(J^m_{P,p^*})^k$, and the entries of $(J^m_{P,p^*})^k$ sum over round trips through all dates—the wage–price spiral, in which prices move desired prices at every date, through nominal costs $mc = mc^r + P$ with weight ω and directly with weight $1-\omega$, and those desired prices move prices again. The amplification factor $\|(I - J^{\text{C}}_{P,p^*}(\phi^{\text{IRF,MC}}))^{-1}\|\,\|(I - J^{\text{MC}}_{P,p^*})^{-1}\|$ in (39) measures the strength of the wage–price spiral: it is large when prices are flexible, and it grows with the reporting horizon, since a permanent movement in real marginal cost feeds inflation indefinitely. At the Phillips-curve level the Calvo approximation is therefore most reliable in sticky-price economies and over moderate horizons.

4. Performance of Robust Statistic Approximation in the Canonical Menu Cost Model

This section presents numerical results on the performance of our method for estimating the degree of monetary non-neutrality in menu cost models.

4.1. Simulation Set-Up

Simulations. From the solution of the stationary problem of the menu cost model, we perform a Monte Carlo simulation to estimate the sufficient statistic. For any given parameterization, the simulation takes as inputs the vector of discretized price gaps x_{grid}, the stationary distribution over price gaps g, the transition matrix of price gaps conditional on no adjustment Q_{not}, the adjustment-probability function $q(x)$, and the target price gap x^*. It then simulates a large cross section of N firms drawn from g, generates a sequence of latent desired prices p^*_{it} using Gaussian shocks ε_{it}, and assembles firm-level pre-adjustment gaps $x^{\text{pre}}_{it} = p_{i,t-1} - p^*_{it}$. Finally, it simulates adjustment from Bernoulli draws according to $q(x^{\text{pre}}_{it})$: if the firm adjusts, it sets $p_{it} = x^* + p^*_{it}$; if it does not adjust, it keeps $p_{it} = p_{i,t-1}$.

With this simulation at hand, we first discard a burn-in period, and then compute the statistic defined in Section 3.1: we estimate the $h \in \{0, 1\}$ local projection (14) of price changes on marginal-cost shocks, infer the synthetic survival rate ϕ^{IRF} using (15), and compute the robust-statistic Calvo Jacobian $J^{\text{IRF}}_{\pi,mc^r}$ as (17).

In addition, from the solution of the model, we compute the true generalized Phillips curve J^{MC}_{π,mc^r}.

Parameterizations. The CalvoPlus model is very flexible. For a given arrival rate of free adjustments, letting the menu cost go to infinity recovers the Calvo model, while keeping a finite menu cost and setting the rate of free adjustments to zero recovers a standard menu cost model. The literature has considered different parameterizations of this general model, with two widely-used calibrations in Golosov and Lucas (2007) and Nakamura and Steinsson (2010). We fix a number of parameters, as detailed in Table 1. We replicate the model-based moments targeted by these two papers, conditional on the parameter values fixed in Table 1, and report the resulting calibrations of the other parameters in Table 2.

Beyond these two standard parametrizations, we show the performance of our approach across a wide range of calibrations. We solve 113 menu cost calibrations that span the mean menu cost, idiosyncratic marginal-cost volatility, and free-adjustment probability. The mean menu cost ranges logarithmically from 0.002 to 0.08, σ ranges linearly from 0.02 to 0.12, and λ ranges linearly from 0 to 0.30, with five values for each parameter. All other parameters are held at their benchmark values. We only keep parameterizations where the quarterly frequency of price changes is at least 14%, keeping 113 out of 125

TABLE 1. Key parameters of the menu cost model

Parameter	notation	value
Discount factor	β	0.99
Elasticity of substitution across varieties	θ	6
Returns to scale	γ	1
Std. of idiosyncratic shocks	σ_ϑ	calibrated
Std. of menu cost	σ_k	0.05
Frequency of free price adjustments	λ	calibrated
Average menu cost	μ_k	calibrated

TABLE 2. Alternative parameterizations of the pricing frictions.

Parameter	Golosov and Lucas	Nakamura and Steinsson
μ_k	0.006	0.051
λ	0.000	0.179
σ_ϑ	0.046	0.06

possible combinations.

Note that because we consider Gaussian shocks as a baseline, the distinction in equations (8)-(9) between the Gaussian identifying shock ϑ and the potentially non-Gaussian unobserved shock z loses content, and we keep ϑ only.

Benchmarks. We define three distinct metrics to assess the quality of our numerical approximation of the generalized Phillips curve. First, we consider the first entry of the Phillips curve Jacobian, which is sufficient to capture the inflationary effects of purely transitory marginal cost changes. Second, we consider distance measures across selected columns of the Phillips curve Jacobian. Third, we consider the Barnichon and Mesters (2021) Phillips multiplier statistic—the ratio between the impulse response functions of inflation and output. This last metric requires to solve the model in general equilibrium, and hence to specify an IS curve and a monetary policy rule. We use standard formulations, detailed in Appendix A.3.

Across these metrics, we compare our robust statistic approach to two benchmarks. First, we consider the corresponding metrics in our IRF statistic generalized Phillips curve assembled from the IRF-implied Calvo hazard $J^{\text{IRF}}_{\pi,mc^r} = J^{\text{C}}_{\pi,mc^r}(\phi^{\text{IRF}})$ compared to the true generalized Phillips curve J^{MC}_{π,mc^r}. Second, we compute the Calvo hazard that best approximates the entire true Jacobian, as defined by Auclert et al. (2024), which we denote ϕ^{ARRS}, and the implied generalized Phillips curve $J^{\text{ARRS}}_{\pi,mc^r} = J^{\text{C}}_{\pi,mc^r}(\phi^{\text{ARRS}})$. Note that computing

ϕ^{ARRS} requires knowing the entire Phillips curve Jacobian, i.e., knowing the structure of the model.

4.2. Results

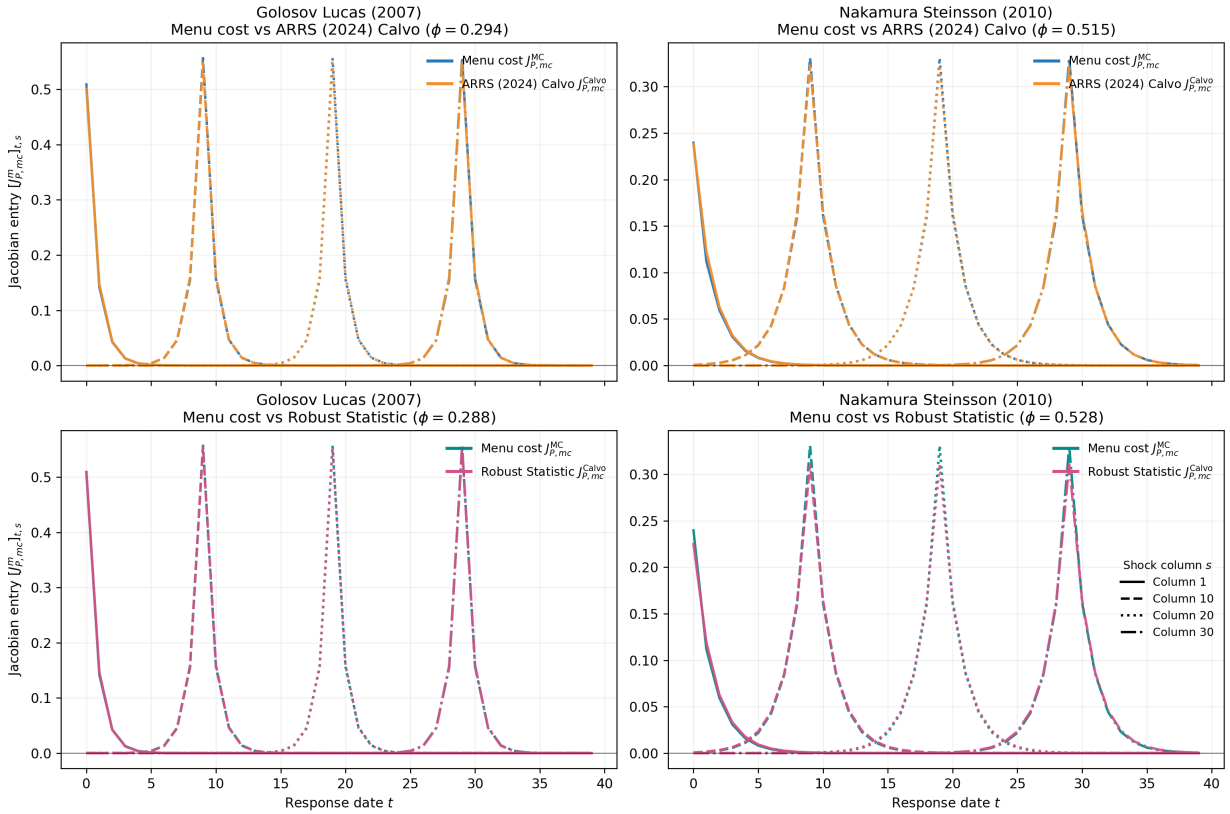
Approximating the menu cost generalized Phillips curve. Figure 1 presents the columns of the true nominal price Jacobian $J^{\text{MC}}_{P,mc}$ and its corresponding approximations $J^{\text{C}}_{P,mc}(\phi)$ for our statistic ϕ^{IRF} and the ARRS approximation ϕ^{ARRS}. The figure shows selected columns of the Jacobian (1st, 10th, 20th, and 30th). The series corresponding to the n-th column of the Jacobian shows the pass-through into the price level in the period in the x-axis of a shock that increases nominal marginal costs in period n. The left plot shows the results for the Golosov-Lucas parameterization while the right plot for the Nakamura-Steinsson parameterization. The top panel shows the results when the Calvo approximation uses the hazard rate from the full Jacobian approximation, the bottom plot uses the hazard rate from the robust statistic.

Both approaches yield remarkable performance across parameterizations. The top panel is a result in Auclert et al. (2024), and the new result we offer in this section is that because the robust statistic infers a hazard rate that is numerically close to the full Jacobian approximation both in Golosov Lucas ($\phi = 0.294$ for ARRS vs. $\phi = 0.288$ for the robust statistic) and in Nakamura Steinsson ($\phi = 0.515$ for ARRS vs. $\phi = 0.528$ for the robust statistic), then the robust statistic approximates the columns of the Jacobian correctly as well. The lower informational burden of the sufficient statistic approach comes at a low cost of small deviations across different columns of the Jacobian for both approaches.

Since the Phillips curve Jacobian transforms the macro pass-through Jacobian using linear algebra, discrepancies in the Phillips curve Jacobian are embedded in the former, but may get amplified or dampened. Figure 2 shows that the structure of the Phillips curve Jacobian is close in both approaches, but the performance is better for Golosov-Lucas, which as ARRS note, paradoxically is closer to a Calvo model than the Nakamura-Steinsson parameterization of the CalvoPlus model.

Beyond the standard parameterizations discussed above, our approach more generally approximates the true Phillips curve Jacobian. We start by showing the performance of the ARRS best Calvo and the sufficient statistic to replicate the first entry of the true Jacobian across calibrations. The first entry of the Jacobian is interesting in its own right because it

FIGURE 1. Nominal price pass-through: selected Jacobian columns

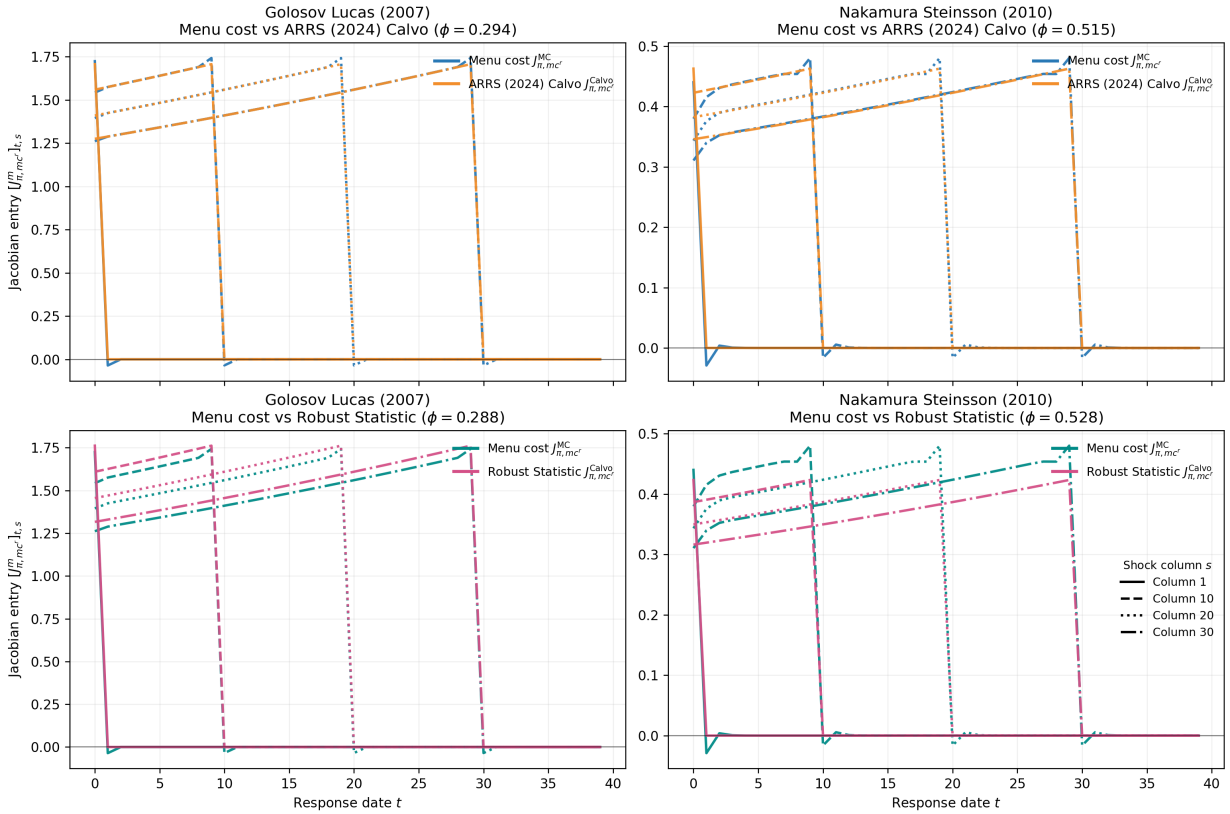


Note: The panels compare columns of the menu-cost nominal pass-through Jacobian $J^{\text{MC}}_{P,mc}$ with their synthetic Calvo counterparts. The left column uses the Golosov–Lucas (2007) calibration and the right column the Nakamura–Steinsson (2010) calibration. Each curve gives the response of the aggregate log price level at date t to a unit nominal marginal-cost disturbance at the indicated date $s \in \{1, 10, 20, 30\}$. Line styles distinguish shock dates. The nominal pass-through Jacobian is constructed from the ARRS best Calvo parameter ϕ^{ARRS} in the top row, and from the IRF statistic ϕ^{IRF} in the bottom row.

is sufficient for the impact inflation response under purely transitory marginal costs. Panel (A) of Figure 3 shows the results. The first entry of the Jacobian is numerically accurate across parameterizations. The slope between the true first entry of the Jacobian and that assembled from the sufficient statistic is 1.003, and the R^2 of the relation between the two objects is 0.999. The underlying DGPs used for these statements have a wide range of inflation responses to a purely transitory shocks, from 0.06 to 100. Panel (B) adds to the figure the behavior of the first entry of the Jacobian using the ARRS Calvo approximation. Both approaches deliver similar approximation accuracy.

To compare performance across models beyond purely transitory shocks, we use the approach of Barnichon and Mesters (2021), who advocate for using the ratio of the present discounted value of the inflation IRF to the present discounted value of the output IRF,

FIGURE 2. Generalized Phillips curve: selected Jacobian columns



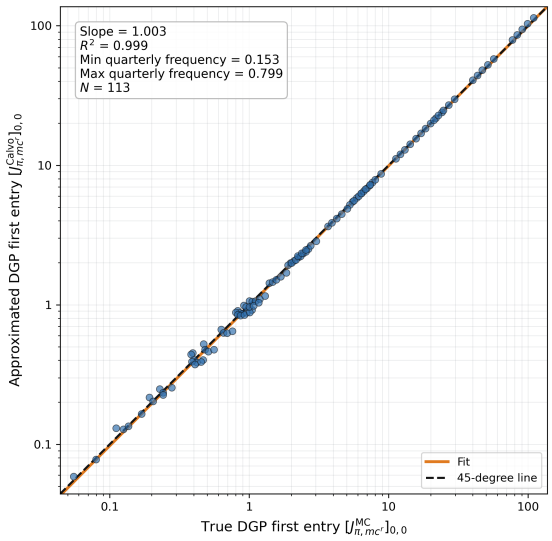
Note: The panels compare columns of J^{MC}_{π,mc^r}, the true generalized Phillips curve Jacobian, with columns of its synthetic Calvo approximations. The left column uses the Golosov–Lucas (2007) calibration and the right column the Nakamura–Steinsson (2010) calibration. The horizontal axis is response date t; line styles identify the date of the unit cost disturbance, $s \in \{1, 10, 20, 30\}$. The generalized Phillips curve is constructed from the ARRS best Calvo parameter ϕ^{ARRS} in the top row, and from the IRF statistic ϕ^{IRF} in the bottom row.

which they call the Phillips multiplier. This normalization is useful because it isolates the properties of the supply block of the model in generating inflationary pressures after a demand shock induces movements in output.

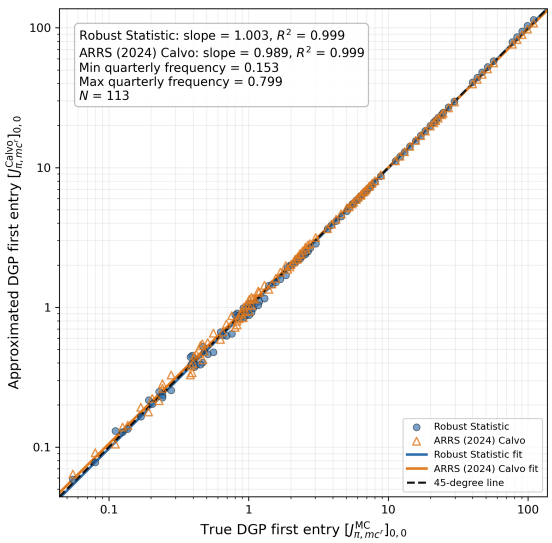
Table 3 reports the Phillips multiplier for the two standard parameterizations of the menu cost model, using the true generalized Phillips curve Jacobian from the model. It also reports the same ratio for the robust-statistic Calvo approximation, and for the best-Calvo approximation from ARRS. The robust statistic performs very well. It matches the ratio of IRFs in the true model almost as well as a Calvo model chosen to approximate the whole Jacobian, without requiring full knowledge of the data-generating process. This result holds more broadly across calibrations: Figure 4(A) shows that the robust statistic approach recovers the correct Phillips multiplier across a wide range of parameterizations. The slope between the true and approximated ratios is 1.008 and an $R^2 = 0.996$. Panel (B)

FIGURE 3. First Entry of the Generalized Phillips Curve

(A) Robust Statistic vs. True Jacobian



(B) Adding the ARRS Best Calvo



Note: Both panels compare the impact entry of an approximating Calvo Phillips curve (vertical axis) with the true generalized Phillips curve entry $J^{\text{MC}}_{\pi,mc^r}[0,0]$ (horizontal axis). Panel A reports the approximation inferred from our robust statistic approach $J^{\text{IRF}}_{\pi,mc^r}[0,0]$. Panel B adds the ARRS best Calvo benchmark $J^{\text{ARRS}}_{\pi,mc^r}[0,0]$. Blue circles denote the $J^{\text{IRF}}_{\pi,mc^r}[0,0]$ approximation; orange triangles in panel (B) denote $J^{\text{ARRS}}_{\pi,mc^r}[0,0]$. Each panel uses the same 113 CES calibrations with quarterly adjustment frequencies above 14% (observed range 15.3–79.9%). Both axes are logarithmic. Solid lines are fitted regressions in logs, and dashed lines denote exact agreement.

adds the same ratio using the ARRS full Jacobian approximation. The results are again very similar.

TABLE 3. Phillips multiplier

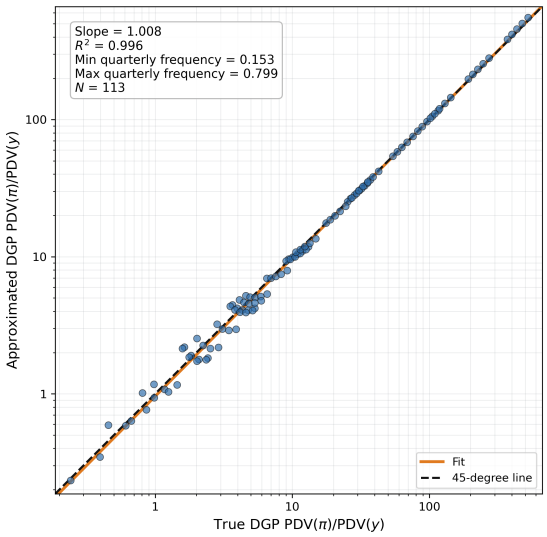
Calibration	True model	Robust statistic $\phi^{\text{IRF,MC}}$	ARRS Best Calvo ϕ^{ARRS}
GL	8.1640	8.4422	8.1813
NS	2.0895	1.9618	2.1437

Note: This table shows the Phillips multiplier, computed as the ratio $\text{PDV}(\Pi)/\text{PDV}(Y)$, for the standard GL and NS calibrations. Discounted sums use a 60-quarter horizon. Both approximations are evaluated using the output sequence from the corresponding menu-cost model.

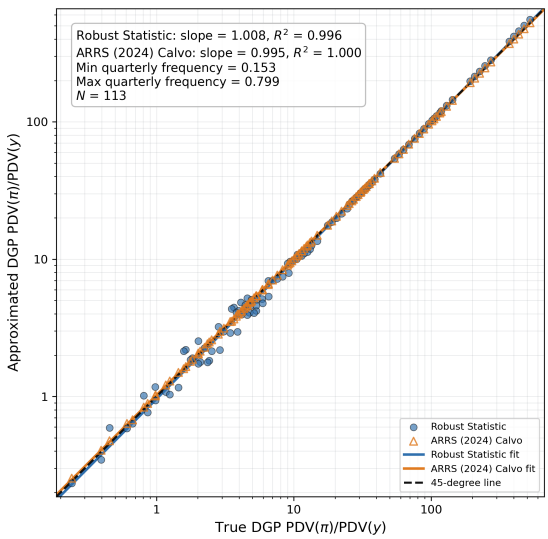
Implied survival rates. A natural skepticism around the results so far is that our robust statistic is accurate because the calibrations we use imply the model behaves very much like a Calvo model. This is not the case. To illustrate that, we plot the "synthetic Calvo" hazard rate inferred from our sufficient statistic and compare it with the alternative of

FIGURE 4. Phillips Multiplier

(A) Robust Statistic vs. True Jacobian



(B) Adding the ARRS Best Calvo



Note: This figure compares the Phillips multiplier $(\text{PDV}(\pi)/\text{PDV}(y))$ computed in a model with the true generalized Phillips curve J^{MC}_{π,mc^r} (horizontal axis) with the one computed in a model with an approximating Calvo Phillips curve (vertical axis). Panel A reports the approximation inferred from our robust statistic approach $J^{\text{IRF}}_{\pi,mc^r}[0,0]$. Panel B adds the ARRS best Calvo benchmark $J^{\text{ARRS}}_{\pi,mc^r}[0,0]$. Blue circles denote the $J^{\text{IRF}}_{\pi,mc^r}[0,0]$ approximation; orange triangles in panel (B) denote $J^{\text{ARRS}}_{\pi,mc^r}[0,0]$. The monetary experiment and the remaining general equilibrium equations are held fixed. Each panel uses the same 113 CES calibrations with quarterly adjustment frequencies above 14% (observed range 15.3–79.9%). Both axes are logarithmic. Solid lines are fitted regressions in logs, and dashed lines denote exact agreement.

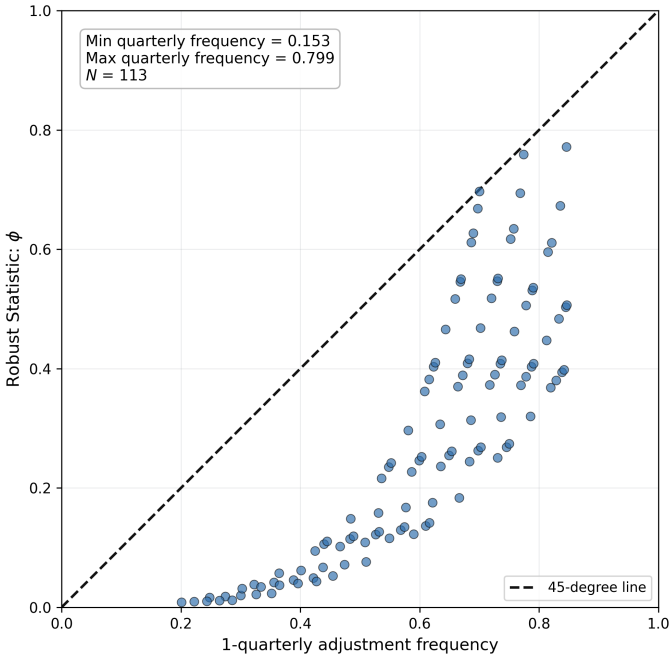
using an unconditional moment and simply computing

$1 -$ frequency of price changes.

The results are presented in Figure 5. If the models were close to Calvo, the robust statistic would recover the frequency of price changes exactly, and all points would lie on top of the 45 degree line. Instead, all points fall below the 45 degree line: the extent of price rigidity that rationalizes the Jacobian is lower than implied by the frequency of price changes. This is because the selection effect implies that, for a given mass of adjusters, prices are more flexible under the menu cost model than under the Calvo model. Some parameterizations are close to the diagonal since the CalvoPlus model nests the Calvo model. The vertical distance between a particular point and the 45 degree line is a model-free measure of the selection effect.

In sum, the goodness of fit of our robust statistic is not by construction, but rather a reflection of the information contained in the dynamic passthroughs.

FIGURE 5. Synthetic Calvo Persistence and the Frequency of Price Changes



Note: The vertical axis reports the synthetic Calvo non-adjustment parameter ϕ^{IRF}, inferred from the first two micro IRF responses b_0 and b_1. The horizontal axis reports the frequency-based Calvo value, $1 - \text{freq}$, where freq is the unconditional quarterly frequency of price adjustment. Each point is one of 113 CES menu-cost calibrations with freq > 0.14. The dashed line is the Calvo benchmark, under which the two quantities coincide.

Robustness to micro-level real rigidities. Our baseline calibration follows GL and NS and assumes constant returns to scale, $\gamma = 1$. We now show that our robust statistic approach performs equally well with micro-level real rigidities. Figure 6 shows the performance of our robust statistic for approximating the Phillips multiplier (panel A) or the first entry of the generalized Phillips curve (panel B), for values of $1 - \gamma$ ranging from 0 (constant returns) to 0.3 (returns to scale equal to 0.7). Our robust statistic tracks the true object extremely well across calibrations of returns to scale.

To highlight the value of starting from conditional moments, we compare the performance of our robust statistic estimate with an approach based on the sufficient statistic result of Alvarez, Le Bihan, and Lippi (2016), as operationalized by Auclert et al. (2024). Because the Calvo model provides a close approximation to the menu cost model, it must approximately have the same cumulative impulse response of output, and hence, the same

ratio of kurtosis to frequency,

$$\frac{\text{Kur}^{\text{C}}}{\text{freq}^{\text{C}}} = \frac{\text{Kur}^{\text{MC}}}{\text{freq}^{\text{MC}}}. \tag{40}$$

In addition, in the discrete-time Calvo model, $\text{Kur}^{\text{C}} = 3(2-\text{freq}^{\text{C}})$. Hence, upon measuring the true ratio $\frac{\text{Kur}^{\text{MC}}}{\text{freq}^{\text{MC}}}$ in the data, one can infer a synthetic Calvo survival probability ϕ^{ABL} that would yield this ratio in a Calvo world

$$\phi^{\text{ABL}} = 1 - \frac{1}{\frac{1}{6}\frac{\text{Kur}^{\text{MC}}}{\text{freq}^{\text{MC}}} + \frac{1}{2}}, \tag{41}$$

and the associated Phillips curve Jacobian $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{ABL}})$.

This alternative method exactly coincides with our method for constant returns to scale, as shown in the dotted green lines of Figure 6. However, deviating from constant returns to scale, our robust statistic tracks the true object, while this alternative statistic deviates from it. This is because moments from the unconditional distribution of price changes cannot tell apart strong real rigidity with large cost shocks and weak real rigidity with small cost shocks, while they have different implications for the Phillips curve. The existing sufficient statistics solve this identification problem by making structural assumptions, like constant returns and Gaussian shocks, with the risk for model mis-specification, as exemplified here.

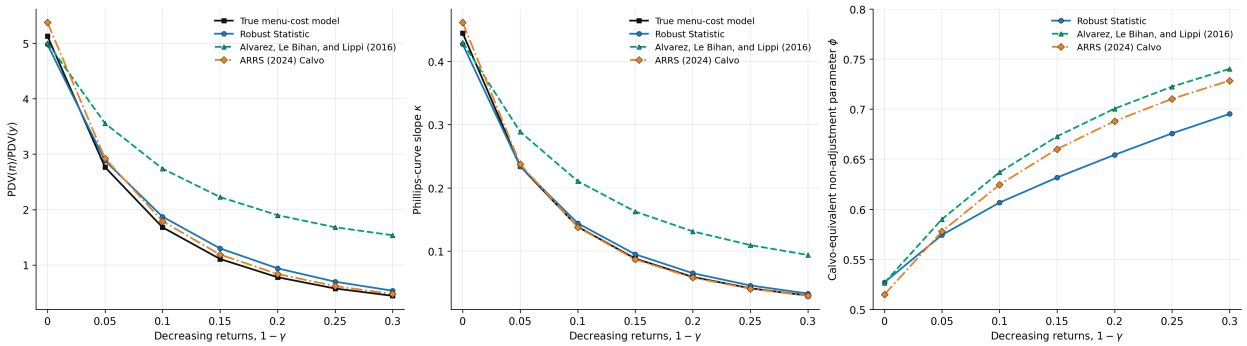
5. Extensions

5.1. Beyond CES: The Kimball Extension

Strategic complementarities change the extent to which firms want to change their prices in the flexible price equilibrium, the incentives that firms have to pay the menu costs, altering the extensive and intensive margins of adjustment, and the size of money non-neutralities. In this section, we document that non-isoelastic demand curves microfounded by a Kimball demand system does not affect the performance of the robust statistic.

As we have shown before, when the true DGP is Calvo, the robust statistic is exact and inclusive of accounting for micro-level real rigidities since the pass-through at every horizon is multiplicative on the flexible price equilibrium desired pass-through ω. Therefore,

FIGURE 6. Performance of the Robust Statistic under Decreasing Returns to Scale



Notes: Parameters remain fixed at the Nakamura Steinsson parameterization. From left to right, the panels report the inflation–output PDV ratio, the first entry of the Phillips curve Jacobian κ, and the Calvo non-adjustment probability ϕ. We compare the model objects obtained using the true generalized Phillips curve J^{MC}_{π,mc^r}, and the Calvo approximations using our robust statistic $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{IRF}})$, the ARRS best Calvo approximation $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{ARRS}})$, and the Calvo survival computed from the Alvarez, Le Bihan, and Lippi (2016) sufficient statistic $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{ABL}})$.

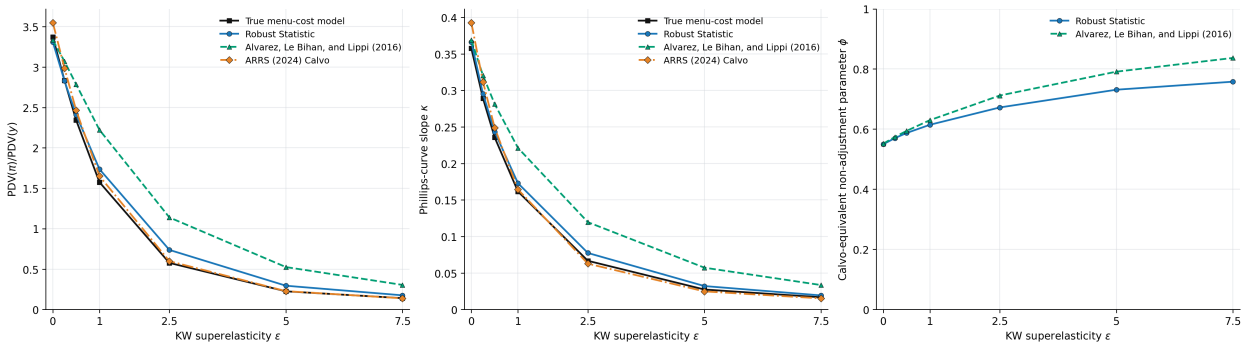
with strategic complementarities, the formula for ω changes and depends on the superelasticity of demand, but the insight that two entries of the micro IRF suffice to identify the generalized Phillips curve remain.

Figure 7 shows the performance of our approximation when varying the superelasticity of demand away from the CES case. The other parameters are fixed at the Nakamura and Steinsson (2010) parameterization of the menu cost model. The zero superelasticity corresponds to the CES case we have discussed already. Panel A shows the behavior of the inflation-output ratio for the True DGP, the robust statistic, and the Calvo approximation using the kurtosis sufficient statistic. Panel (B) shows the slope of the Phillips curve that the three Calvo approximations infer. Panel (C) shows the synthetic non-adjustment probability inferred by the three approximation methods. The robust statistic tracks the slope of the Phillips curve and the macro inflation dynamics of the best Calvo approximation and the true DGP, respectively. Even though the kurtosis statistic tracks the inferred ϕ closely, it then infers that the extent of monetary non-neutrality is smaller. For instance, when the superelasticity is equal to 2.5, a 1 percentage point increase in the present discounted value of output generates a 0.5 percentage point increase in inflation, while the kurtosis statistic infers it would generate more than twice the inflationary pressure.

In Figure C.1 we show the performance of the robust statistic versus the ARRS best Calvo approximation of the true Jacobian for a wide range of parameterizations. This fit

is unchanged by the value of the superelasticity going from zero (CES) to 7.5. Panel B compares the ARRS best Calvo approximation to the approximation inferred from the Alvarez, Le Bihan, and Lippi (2016) sufficient statistic (equation 41). The average bias of the kurtosis-to-frequency statistic increases monotonically with the superelasticity of demand, and to large values, when the departures with respect to CES are economically meaningful.

FIGURE 7. Performance of the Robust Statistic under Non-Isoelastic Demand Curves



Notes: The figure holds the baseline economic parameters fixed and varies the Klenow–Willis demand superelasticity $\varepsilon \in \{0, 0.25, 0.5, 1, 2.5, 5, 7.5\}$. Idiosyncratic cost persistence is $\rho = 0.95$. The left panel compares the discounted inflation-to-output ratio, $\text{PDV}(\pi)/\text{PDV}(y)$, in the true menu-cost model with the ratios implied by the robust micro statistic and the kurtosis-to-frequency statistic. The middle panel compares the inferred Phillips-curve slopes with the slope of the ARRS (2024) best-fitting Calvo approximation. The right panel reports the Calvo-equivalent nonadjustment parameter ϕ inferred by the two statistics. Blue circles denote the robust statistic, and green triangles denote the Alvarez, Le Bihan, and Lippi statistic.

Figure C.2 shows the performance of our approximation of the Phillips multiplier when varying the superelasticity of demand. The macro dynamics assembled from the robust statistic are well replicated both in CES and away from the CES limit.

5.2. Measurement Error

Measurement error in prices is a challenge for the measurement of monetary non-neutrality starting from unconditional micro price data. First, measurement error biases frequency of price change measures towards one and affects higher moments of price changes as well, crucially the kurtosis of price changes. Under classical Gaussian measurement error that is sufficiently large with respect to the standard deviation of true prices, the measured frequency of price changes goes to 1, and due to normality in measurement error, the kurtosis goes to 3. The question we answer in this section is the comparative behavior of our sufficient statistic and the Alvarez, Le Bihan, and Lippi (2016) statistic under measurement

error.

We introduce measurement error as an intermediate step between model simulated data and the data that an econometrician observes. In particular, an econometrician observes $p_{it}^{obs} = p_{it} + me_{it}$, where the measurement error term me_{it} is i.i.d and Gaussian with standard deviation σ_{me}. Since the moments the literature uses are about price changes, then the relevant volatility of the measurement error is that of Δme_{it} which is equal to $\sqrt{2}\sigma_{me}$. To ease the interpretation of measurement error we present it as a ratio of the volatility of productivity shocks ϑ_{it}, which we denote σ_{ϑ}. For a given panel of prices $\{p_{it}\}_{\forall i,t}$, we generate a large number of samples of measurement error of varying volatility, from values ranging $\sqrt{2}\sigma_{me}/\sigma_{\vartheta} \in [0, 50\%]$. We generate 10%-90% confidence bands from the set of samples to generate confidence intervals. We follow the guideline of excluding from the computation of the kurtosis-to-frequency statistic price changes (genuine or not) smaller than 0.1% in absolute value.

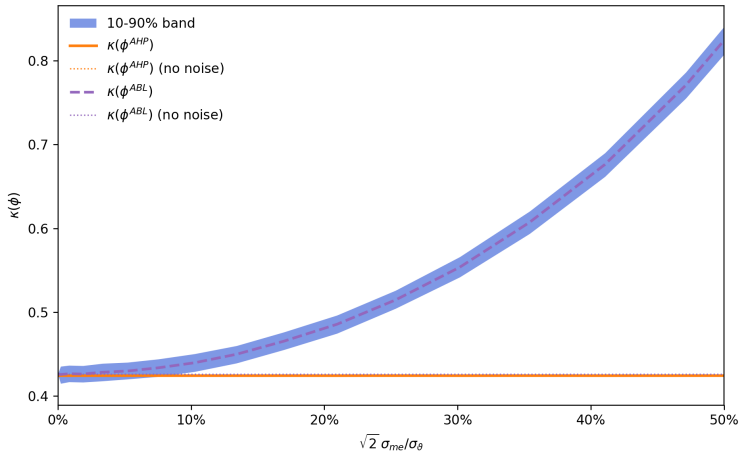
Figure 8 shows that measurement error does not bias the inferred slope of the Phillips curve from the sufficient statistic $J^{\text{IRF}}_{\pi,mc^r}$, or the inferred Calvo survival rate ϕ^{IRF}. The robust statistic is unaffected by measurement error, an advantaged caused by the noisy variable being the dependent variable in our regressions. Measurement error increases the noise in the robust statistic, although that is not quantitatively important (the error bands in the robust statistic are not visible in the plot).

By contrast, the effects of measurement error on the kurtosis-to-frequency statistic, and hence on the inferred Phillips curve $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{ABL}})$, are very different. Initially, measurement error increases both kurtosis (there is more mass in small price changes) and frequency which roughly offset. Little by little, measurement error decreases the kurtosis while its effects on the frequency are monotonic. Through the lens of the sufficient statistic that means that monetary non-neutrality decreases, which can only be rationalized in a Calvo model with a steeper slope of the Phillips curve.

5.3. Non-Random Walk Shocks

Random-walk shocks are appealing due to their tractability. A productivity shock that changes the price gap today will permanently change future expected price gaps in the absence of adjustment. Transitory shocks are different; a firm will consider the persistence of the shock when deciding whether to act on it, changing the size of the inaction region.

FIGURE 8. Price Measurement Error and Inferred Phillips Curve Slopes

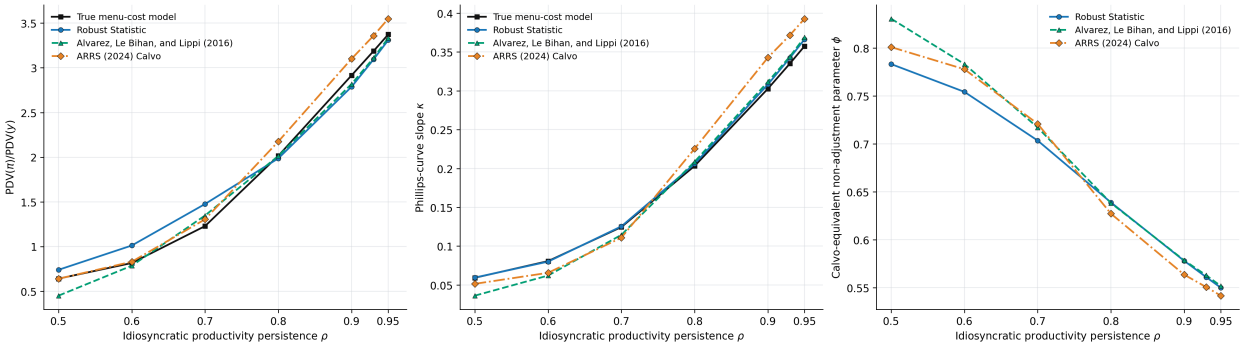


Notes: Parameters are fixed at the Nakamura–Steinsson calibration. Observed log prices follow $p_{it}^{\text{obs}} = p_{it} + me_{it}$, with independent Gaussian measurement error in price levels. The horizontal axis is the standard deviation of measurement error in price changes relative to the cost-innovation standard deviation, $\text{sd}(\Delta me)/\sigma_\vartheta = \sqrt{2}\sigma_{me}/\sigma_\vartheta$. The solid orange line is the replication mean of the slope inferred from our robust statistic approach $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{IRF}})$ (denoted AHP in the legend). The dashed purple line is the mean slope inferred from the Alvarez, Le Bihan, and Lippi (2016) statistic $J^{\text{C}}_{\pi,mc^r}(\phi^{\text{ABL}})$. Shading shows the 10th–90th percentile range across replications; dotted lines show each method's no-noise value. The ABL calculation counts observed absolute log price changes above 0.001 when computing frequency and kurtosis.

Sufficient statistics in the literature use the random walk assumption, but are not reliable when the DGP features other assumptions of shocks (Dotsey and King 2005).

Our robust statistic works well regardless of the persistence of idiosyncratic shocks. We show that numerically by solving menu cost models with AR(1) productivity shocks with varying persistence. Figure 9 shows that the robust statistic tracks the true Phillips multiplier of the menu cost model (left panel), the slope of the Phillips curve of the best Calvo approximation of the full Jacobian (middle panel), and the virtual frequency of price changes.

FIGURE 9. Performance of the Robust Statistic Across Shock Persistence



Notes: The figure varies the persistence of idiosyncratic productivity shocks under CES demand, keeping the remaining Nakamura Steinsson benchmark primitives fixed. Log physical productivity follows $\vartheta_{it} = \rho\vartheta_{i,t-1} + \epsilon_{it}$, with $\rho \in \{0.50, 0.60, 0.70, 0.80, 0.90, 0.93, 0.95\}$. The innovation standard deviation remains fixed, hence stationary productivity dispersion changes with ρ. From left to right, the panels report the inflation–output PDV ratio, the Phillips-curve slope κ, and the Calvo-equivalent non-adjustment probability α. The left panel compares the true menu-cost model with the Robust Statistic approximation, evaluated on the same true-model marginal-cost and output paths. The middle and right panels compare the Robust Statistic with the ARRS (2024) Calvo approximation to the menu-cost Phillips-curve Jacobian.

6. Conclusion

In this paper, we propose a robust statistic for monetary non-neutrality: the impact and one-period-ahead responses of a firm's price to an observed idiosyncratic cost shock, estimated by a local projection and assembled into a Phillips curve as if the economy were Calvo.

Three results support this statistic. First, in the Calvo, Gertler–Leahy, and canonical menu cost models, the average response of firms' prices to their own cost shock equals the response of the aggregate price level to a permanent nominal cost shock, so that a firm-level regression measures the object the Phillips curve is built from. Second, the two

coefficients recover the generalized Phillips curve exactly in the Calvo and Gertler–Leahy models, where the decay of a cost shock in firms' price gaps is geometric. Third, in menu cost models, the two coefficients recover the average one-period-ahead survival rate of a marginal price gap and extrapolating this survival geometrically is accurate across a wide range of parameterizations. The accuracy survives decreasing returns to scale, Kimball demand, measurement error in prices, and persistent rather than random-walk shocks.

Our approach contrasts with sufficient statistics based on unconditional moments of the distribution of price changes, which tend to quickly lose informativeness as soon as the true data-generating process deviates from the benchmark model under which these statistics are sufficient.

References

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi. 2016. "The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach." *American Economic Review* 106 (10): 2817–2851.

Alvarez, Fernando, and Francesco Lippi. 2014. "Price Setting with Menu Cost for Multiproduct Firms." *Econometrica* 82 (1): 89–135.

Alvarez, Fernando, Francesco Lippi, and Aleksei Oskolkov. 2022. "The Macroeconomics of Sticky Prices with Generalized Hazard Functions." *The Quarterly Journal of Economics* 137 (2): 989–1038.

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis. 2023. "Price Setting with Strategic Complementarities as a Mean Field Game." *Econometrica* 91 (6): 2005–2039.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2019. "International Shocks, Variable Markups, and Domestic Prices." *Review of Economic Studies* 86 (6): 2356–2402.

Angrist, Joshua D., and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, vol. 3A, 1277–1366. Amsterdam: Elsevier.

Auclert, Adrien, Rodolfo D. Rigato, Matthew Rognlie, and Ludwig Straub. 2024. "New Pricing Models, Same Old Phillips Curves?" *The Quarterly Journal of Economics* 139 (1): 121–186.

Ball, Laurence, and David Romer. 1990. "Real Rigidities and the Non-Neutrality of Money." *The Review of Economic Studies* 57 (2): 183–203.

Barnichon, Régis, and Geert Mesters. 2021. "The Phillips Multiplier." *Journal of Monetary Economics* 117: 689–705.

Caballero, Ricardo J., and Eduardo M. R. A. Engel. 2007. "Price Stickiness in Ss Models: New Interpretations of Old Results." *Journal of Monetary Economics* 54 (Supplement 1): 100–121.

Calvo, Guillermo A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12 (3): 383–398.

Caplin, Andrew S., and Daniel F. Spulber. 1987. "Menu Costs and the Neutrality of Money." *The Quarterly Journal of Economics* 102 (4): 703–726.

Cerrato, Andrea, and Giulia Gitti. 2022. "Inflation since covid: Demand or supply."

Dotsey, Michael, and Robert G. King. 2005. "Implications of State-Dependent Pricing for Dynamic Macroeconomic Models." *Journal of Monetary Economics* 52 (1): 213–242.

Dotsey, Michael, and Alexander L. Wolman. 2020. "Investigating Nonneutrality in a State-Dependent Pricing Model with Firm-Level Productivity Shocks." *International Economic Review* 61 (1): 159–188.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo. 2011. "Reference Prices, Costs, and Nominal Rigidities." *American Economic Review* 101 (1): 234–262.

Fitzgerald, Terry, Callum Jones, Mariano Kulish, and Juan Pablo Nicolini. 2024. "Is there a stable relationship between unemployment and future inflation?" *American Economic Journal: Macroeconomics* 16 (4): 114–142.

Gagliardone, Luca, Mark Gertler, Simone Lenzu, and Joris Tielens. 2025. "Anatomy of the Phillips Curve: Micro Evidence and Macro Implications." *American Economic Review* 115 (11): 3941–3974.

Gertler, Mark, and John Leahy. 2008. "A Phillips Curve with an Ss Foundation." *Journal of Political Economy* 116 (3): 533–572.

Golosov, Mikhail, and Robert E. Lucas, Jr. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115 (2): 171–199.

Gopinath, Gita, and Oleg Itskhoki. 2010. "Frequency of Price Adjustment and Pass-Through." *The Quarterly Journal of Economics* 125 (2): 675–727.

Härdle, Wolfgang, and Thomas M. Stoker. 1989. "Investigating Smooth Multiple Regression by the Method of Average Derivatives." *Journal of the American Statistical Association* 84 (408): 986–995.

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson. 2022. "The Slope of the Phillips Curve: Evidence from U.S. States." *The Quarterly Journal of Economics* 137 (3): 1299–1344.

Herreño, Juan, Noémie Pinardon-Touati, and Malte Thie. 2026. "How Steep is the Phillips Curve in Developing Economies? A Sufficient Statistics Approach and Estimates for India." Working paper, March 2026.

Hong, Gee Hee, Matthew Klepacz, Ernesto Pastén, and Raphael Schoenle. 2023. "The Real Effects of Monetary Shocks: Evidence from Micro Pricing Moments." *Journal of Monetary Economics* 139: 1–20.

Karadi, Peter, and Adam Reiff. 2019. "Menu Costs, Aggregate Fluctuations, and Large Shocks." *American Economic Journal: Macroeconomics* 11 (3): 111–146.

Kehoe, Patrick, and Virgiliu Midrigan. 2015. "Prices Are Sticky After All." *Journal of Monetary Economics* 75: 35–53.

Kimball, Miles S. 1995. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit and Banking* 27 (4): 1241–1277.

Klenow, Peter J., and Jonathan L. Willis. 2016. "Real Rigidities and Nominal Price Changes." *Economica* 83 (331): 443–472.

Midrigan, Virgiliu. 2011. "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations." *Econometrica* 79 (4): 1139–1180.

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

Nakamura, Emi, and Jón Steinsson. 2010. "Monetary Non-Neutrality in a Multisector Menu Cost

Model." *The Quarterly Journal of Economics* 125 (3): 961–1013.

Rotemberg, Julio J. 1982. "Monopolistic Price Adjustment and Aggregate Output." *The Review of Economic Studies* 49 (4): 517–531.

Sangani, Kunal. 2026. "Complete pass-through in levels." *The Quarterly Journal of Economics* 141 (2): 1077–1135.

Stein, Charles. 1972. "A Bound for the Error in the Normal Approximation to the Distribution of a Sum of Dependent Random Variables." In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, edited by Lucien M. Le Cam, Jerzy Neyman, and Elizabeth L. Scott, 583–602. Berkeley: University of California Press.

Stein, Charles M. 1981. "Estimation of the Mean of a Multivariate Normal Distribution." *The Annals of Statistics* 9 (6): 1135–1151.

Stoker, Thomas M. 1986. "Consistent Estimation of Scaled Coefficients." *Econometrica* 54 (6): 1461–1481.

Yitzhaki, Shlomo. 1996. "On Using Linear Regressions in Welfare Economics." *Journal of Business & Economic Statistics* 14 (4): 478–486.

Appendix

Appendix A. Additional Details of the Model Environments

A.1. The Loss Function from Primitives

Nominal marginal costs of firm i at date t are

$$MC_{it} = \frac{W_t}{\gamma A_{it}} \Big(\frac{Y_{it}}{A_{it}} \Big)^{\frac{1-\gamma}{\gamma}}, \tag{A.1}$$

The desired price. Write the profit (2) as revenue minus cost, $R(P_{it}) = (P_{it}/P_t)^{1-\varepsilon} Y_t$ and $K(P_{it}) = \frac{W_t}{P_t}(Y_t/A_{it})^{1/\gamma}(P_{it}/P_t)^{-\varepsilon/\gamma}$. The first-order condition is

$$K(P_{it}^*) = \frac{\gamma(\varepsilon-1)}{\varepsilon}\, R(P_{it}^*), \tag{A.2}$$

the markup condition $P_{it}^* = \frac{\varepsilon}{\varepsilon-1} MC_{it}$ with $MC_{it} = P_t K'(Y)$ from (A.1). Taking logs and using $\frac{\varepsilon}{\gamma} - (\varepsilon - 1) = 1/\omega$,

$$\ln P_{it}^* = \omega \Big(\ln W_t + \frac{1-\gamma}{\gamma} \ln Y_t - \frac{1}{\gamma} \ln A_{it} \Big) + (1-\omega) \ln P_t + \text{const}, \tag{A.3}$$

which is (3) in log deviations, with $mc_t = w_t + \frac{1-\gamma}{\gamma} Y_t$ the nominal marginal cost of a firm with $A_{it} = 1$ producing Y_t.

Gap sufficiency and the curvature. Substituting $P_{it} = P_{it}^* e^x$ and using (A.2),

$$\Pi_t(P_{it}^* e^x) = R_{it}^*\, h_\omega(x), \qquad h_\omega(x) \equiv e^{(1-\varepsilon)x} - \frac{\gamma(\varepsilon-1)}{\varepsilon}\, e^{-(\varepsilon/\gamma)x}, \tag{A.4}$$

where $R_{it}^* \equiv R(P_{it}^*) = Y_t (P_{it}^*/P_t)^{1-\varepsilon}$ is frictionless revenue. Profit as a function of the log price gap is therefore a fixed shape h_ω, which depends only on (ε, γ), times a scale R_{it}^* that carries all the dependence on the firm's productivity and on the aggregates. This is the property that makes the gap a sufficient state for the firm's pricing problem *exactly*, not only to second order.[19] A second-order expansion of (A.4) around $x = 0$ gives (4), with

[19] It fails under Kimball demand, where the elasticity of demand varies with the relative price so that the shape of profit in the gap depends on the firm's relative cost; there the reduction to a gap problem holds

curvature $C_{it} = (\varepsilon - 1)R^*_{it}/\omega$.

From (4) **to the objective** (5)**.** The curvature factors as $C_{it} = \frac{\varepsilon-1}{\omega} R^*_{it}$, and by (A.3) the desired relative price is

$$\ln \frac{P^*_{it}}{P_t} = \omega\Big(\ln \frac{W_t}{P_t} + \frac{1-\gamma}{\gamma} \ln Y_t - \frac{1}{\gamma} \ln A_{it} \Big) + \text{const},$$

so

$$C_{it} \propto R^*_{it} \propto A_{it}^{\omega(\varepsilon-1)/\gamma} \times \text{aggregate terms}, \tag{A.5}$$

Following Gertler and Leahy (2008), we keep the adjustment problem homogeneous in productivity, andwe take the menu cost to be a fixed fraction of the firm's frictionless revenue: menu cost$_{it} = \xi\, C_{it}$ with ξ a constant. Dividing the date-t objective by C_{it} then gives the per-period loss $\frac{1}{2}x^2_{it}$ and the menu cost ξ, common across firms and dates, for any size of the shocks.

The only place where the curvature survives is the continuation value, in which the ratio $C_{i,t+s}/C_{it} = R^*_{i,t+s}/R^*_{it}$ weights future losses. Here the order of approximation enters, and it is the same one under which (4) was written. Consider the joint perturbation of order υ in which the innovations to a_{it}, mc_t and P_t and the band half-width $\bar{x} \propto \sqrt{\xi}$ are all first order. In that perturbation the gap is $O(\upsilon)$ and the ratio $C_{i,t+s}/C_{it}$ is $1 + O(\upsilon)$, so its variation multiplies $x^2 = O(\upsilon^2)$ and is $O(\upsilon^3)$, the order already dropped in (4).[20]. The household's stochastic discount factor is replaced by β^t for the same reason.

The normalized problem is the canonical menu cost problem: a quadratic loss $\frac{1}{2}x^2_{it}$, a menu cost ξ common to all firms and dates, and i.i.d. innovations $\omega(\epsilon^z_{it} + \epsilon_{it})$ to the desired price. Every property of the canonical problem used in the text holds only up to terms of order υ^3.

only to second order.

[20]This is the term that Gertler and Leahy (2008)'s Proposition 1 shows to be independent of the current price to second order, and that vanishes in Auclert et al. (2024)'s small-risk limit (their Proposition 4)

A.2. The Gertler–Leahy Economy without Islands

This section restates the parts of Gertler and Leahy (2008) needed for Proposition 3 in the notation of Section 2.1.[21]

A.2.1. Assumptions and the Firm's Problem

Primitives. A firm is hit by turbulence with probability $1-\eta$ per period and, independently, draws a free adjustment opportunity with probability λ, both i.i.d.; write $\eta_\lambda \equiv (1-\lambda)\eta$ for the probability that neither occurs. At a hit $z_{it} = z_{i,t-1} + \epsilon^z_{it}$ with $\omega\epsilon^z_{it} \sim U[-\varsigma/2, \varsigma/2]$ (so that ς is the support width of the turbulence shock in desired-price units), otherwise $z_{it} = z_{i,t-1}$. The firm's desired price is (7). Changing a price costs the menu cost ξ (zero at a free draw). Following GL, a firm knows when it is hit but, to learn the value of the draw and to reconsider its price, it must first pay a small decision cost d; a firm that does not pay it keeps its price. A free adjustment opportunity waives both ξ and d: the firm observes its state and resets at no cost. GL's assumptions are:

- (A1) the menu cost ξ is second order, so that the band half-width $\bar{x} = O(\sqrt{\xi})$ is first order;
- (A2) innovations to mc_t, P_t and ϑ_{it} are first-order small: ϵ_{it} has bounded support, and the cumulative movement of the non-turbulence part of the desired price, $\omega(mc_t + \vartheta_{it}) + (1-\omega)P_t$, over any $k(m)$ consecutive periods is bounded by m for every realization, where $k(m)$ is the least integer with $\eta_\lambda^{k(m)} \leq cm$ for a constant $c > 0$;
- (A3) $\varsigma > 4\bar{x} + 2m$;
- (A4) the decision cost d lies in the range for which paying it is optimal after a hit and not after a period with neither a hit nor a free draw.

A.2.2. Characterization of the Optimal Policy

Policy function. By (A4), a firm reconsiders its price only when it is hit or receives a free adjustment opportunity; (A1)–(A3) make the range in (A4) nonempty, since the gain from reconsidering without a hit is of order $m(2\bar{x}+m)$ while the gain after a hit is of order ς^2.

[21] Compared to Gertler and Leahy (2008), (i) we add the free adjustment probability λ of (MC), (ii) we drop the entry–exit process; (iii) the real rigidity is generated by decreasing returns as opposed to island-specific labor markets.

Under (A1)–(A4), GL's Proposition 1 states that the expected value of an optimal policy after the next hit is independent of the firm's current gap to second order (the same holds after a free draw, which resets the gap); so that in choosing its price the firm only takes into account the states in which neither a further hit nor a free draw has occurred, weighted by $(\eta_\lambda\beta)^s$. With the quadratic loss, the value of a firm that has just reset to p at date t is, up to terms independent of p and third-order terms,

$$\text{(A.6)} \qquad \begin{aligned} &-\frac{1}{2}\,\mathbb{E}_t\sum_{s\geq 0}(\eta_\lambda\beta)^s(p-p^*_{i,t+s})^2, \\ &p^*_{i,t+s}=\omega(mc_{t+s}+z_{it}+\vartheta_{i,t+s})+(1-\omega)P_{t+s} \text{ along the no-hit path.} \end{aligned}$$

The first-order condition $\mathbb{E}_t\sum_s(\eta_\lambda\beta)^s(p-p^*_{i,t+s})=0$ gives the reset price

$$\text{(A.7)} \qquad p^{\text{new}}_{it}=\omega z_{it}+(1-\eta_\lambda\beta)\,\mathbb{E}_t\sum_{s\geq 0}(\eta_\lambda\beta)^s[\omega(mc_{t+s}+\vartheta_{i,t+s})+(1-\omega)P_{t+s}],$$

which is (27). Since z_{it} is frozen along the no-hit path it enters with coefficient ω; with a random-walk shifter, $\mathbb{E}_t\vartheta_{i,t+s}=\vartheta_{it}$ and $p^{\text{new}}_{it}=\omega(z_{it}+\vartheta_{it})+(1-\eta_\lambda\beta)\mathbb{E}_t\sum_s(\eta_\lambda\beta)^s[\omega mc_{t+s}+(1-\omega)P_{t+s}]$. In the zero-inflation stationary state mc and P are constant, so $p^{\text{new}}_{it}=p^*_{it}$: the firm resets to its current desired price, and the reset gap is zero.

The inaction band. Write the value (A.6) as a function of the gap y between the firm's price and the reset price (A.7); it equals $-\frac{1}{2}\,y^2/(1-\eta_\lambda\beta)$ plus terms independent of y. Value matching equates the value of keeping a price with gap y to the value of resetting net of the menu cost, $\frac{1}{2}\,y^2/(1-\eta_\lambda\beta)=\xi$, so the firm adjusts if and only if $|y|>\bar{x}$ with

$$\text{(A.8)} \qquad \bar{x}=\sqrt{2(1-\eta_\lambda\beta)\,\xi},$$

an expression analogous to Gertler and Leahy (2008).[22]

Two remarks on what the band is a statement about. First, it is a band on the gap to the *reset* price, $y=p_{i,t-1}-p^{\text{new}}_{it}$, not on the gap $x_{it}=p_{it}-p^*_{it}$ to the flexible-price desired price of Section 2.1: reconsidering means comparing the current price with the price

[22] For $\lambda=0$ and a menu cost b in output units, Gertler and Leahy (2008)'s band is $\bar{x}^2=2(1-\eta\beta)\,b/((\varepsilon-1)\bar{Y})$; here $b=\xi\bar{C}$ with $\bar{C}=(\varepsilon-1)\bar{Y}/\omega$ at the symmetric steady state $\bar{P}^*=\bar{P}$, so $\bar{x}^2=2(1-\eta_\lambda\beta)\,\omega b/((\varepsilon-1)\bar{Y})$: for the same output-unit menu cost the band is narrower than GL's by $\sqrt{\omega}$, because decreasing returns steepen the firm's profit in its own price whereas GL's island labor market does not.

the firm would set, and by GL's Proposition 1 that comparison depends on y alone. The two gaps coincide in the zero-inflation stationary state, where $p_{it}^{\text{new}} = p_{it}^{*}$, and after an unanticipated permanent shift of the desired price with the price level held fixed, since the problem is stationary again from the shock date on. They differ along anticipated aggregate paths, where the reset price is forward-looking. Second, the band is not just a steady-state object: its half-width (A.8) is constant along every aggregate path admissible under (A2), because the coefficient $1/(1-\eta_\lambda\beta)$ of the quadratic value does not depend on the state. What moves with the aggregate path is the center of the band, the reset price. The drift in the reset price between two hits is bounded by m and (A2)–(A3) keep the band strictly inside the support of the turbulence draw along every such path.

A.3. General Equilibrium Structure

We solve and simulate a wide range of New Keynesian models whose linearized aggregate block consists of a monetary policy rule, an intertemporal IS curve and a generalized Phillips curve. Let $\boldsymbol{\pi}$ denote the inflation sequence, $\tilde{\mathbf{y}}$ the sequence of log deviations of output from its flexible-price level, $\mathbf{i}$ the nominal interest rate and $\boldsymbol{\zeta}$ the monetary-policy disturbance. The monetary policy rule is

$$\mathbf{i} = \psi_\pi \, \boldsymbol{\pi} + \boldsymbol{\zeta},$$

with ψ_π the inflation coefficient. The IS curve is

$$\tilde{\mathbf{y}} = -\frac{1}{\sigma_c}\,(I-F)^{-1}(\mathbf{i} - F\boldsymbol{\pi}), \qquad \text{i.e.} \qquad \tilde{y}_t = -\frac{1}{\sigma_c}\sum_{s\geq 0}(i_{t+s} - \pi_{t+s+1}),$$

with σ_c the inverse intertemporal elasticity of substitution, where $(I-F)^{-1}$ is the upper-triangular matrix of ones and $F(I-F)^{-1}$ its strictly upper-triangular part. The generalized Phillips curve is

$$\boldsymbol{\pi} = J^{m}_{\pi,mc^r}\,\mathbf{mc}^r, \qquad \mathbf{mc}^r = \chi\,\tilde{\mathbf{y}}, \qquad \text{so that} \qquad \boldsymbol{\pi} = J^{m}_{\pi,\tilde{y}}\,\tilde{\mathbf{y}}, \quad J^{m}_{\pi,\tilde{y}} \equiv \chi\, J^{m}_{\pi,mc^r},$$

with χ the elasticity of real marginal cost to output. When the Phillips curve is assembled from the IRF statistics, J^{m}_{π,mc^r} is replaced by $J^{\text{IRF}}_{\pi,mc^r} = \kappa^{\text{IRF}}\Theta_\beta$, equation (17).

The monetary-policy disturbance is an AR(1) impulse of size ζ_0 and persistence ρ_ζ,

$$\boldsymbol{\zeta} = \zeta_0 \,(1, \rho_\zeta, \rho_\zeta^2, \dots)',$$

truncated at the horizon of the sequence.

Substituting the rule into the IS curve and the result into the Phillips curve gives a linear system for inflation,

$$\Big[I + \frac{1}{\sigma_c}\, J^m_{\pi,\tilde{y}}\,(I-F)^{-1}(\psi_\pi I - F)\Big]\boldsymbol{\pi} = -\frac{1}{\sigma_c}\, J^m_{\pi,\tilde{y}}\,(I-F)^{-1}\boldsymbol{\zeta},$$

which we solve by inverting the bracket on the truncated system. Given $\boldsymbol{\pi}$, the output and nominal-rate IRFs follow from the IS curve and the monetary policy rule.

Appendix B. Proofs of the Theoretical Results

B.1. Calvo

B.1.1. Proof of Lemma 1

Proof. (i) $\Theta(\alpha)\mathbf{1} = (1-\beta\alpha)^{-1}\mathbf{1}$ and $[\Upsilon(\alpha)\mathbf{1}]_h = \sum_{r=0}^{h}\alpha^r = (1-\alpha^{h+1})/(1-\alpha)$. (ii) Write $c \equiv (1-\alpha)(1-\beta\alpha)$ and $J \equiv J^{\mathrm{C}}_{P,p^*}(\alpha)$. Since $(I-J)^{-1}J = (J^{-1}-I)^{-1}$ and $J^{-1} = c^{-1}\Theta(\alpha)^{-1}\Upsilon(\alpha)^{-1} = c^{-1}(I-\beta\alpha F)(I-\alpha L)$, using $FL = I$,

$$J^{-1} - I = c^{-1}\big[(1+\beta\alpha^2 - c)I - \alpha L - \beta\alpha F\big] = \frac{\alpha}{c}\big[(1+\beta)I - L - \beta F\big] = \frac{\alpha}{c}(I-\beta F)(I-L),$$

where the second equality uses $1+\beta\alpha^2-(1-\alpha)(1-\beta\alpha) = \alpha(1+\beta)$ and the third again uses $FL = I$. Inverting gives $(I-J)^{-1}J = (c/\alpha)(I-L)^{-1}(I-\beta F)^{-1}$, and premultiplying by $(I-L)$ gives $(I-L)(I-J)^{-1}J = (c/\alpha)(I-\beta F)^{-1} = \kappa(\alpha)\Theta_\beta$. □

B.2. Rotemberg

This section adds a third adjustment technology to the environment of Section 2.1, the convex adjustment cost of Rotemberg (1982), and shows that the IRF statistics are exact there as well. Model (R) keeps the firms, the desired price (7) and the idiosyncratic shocks of (C): ϵ_{it} and ϵ^z_{it} can have any distribution with the moments stated in Section 2.1. The

adjustment technology differs. There is no menu cost and no free draw; instead, changing the price by $p_{it}-p_{i,t-1}$ costs $\frac{\nu}{2}(p_{it}-p_{i,t-1})^2$ per period in the curvature units of Appendix A.1, with $\nu>0$ a constant.[23] Every price changes every period, so freq $=1$.

The firm's problem. The firm chooses $\{p_{it}\}$ to minimize

$$\mathbb{E}_0\sum_{t\geq 0}\beta^t\Big[\frac{1}{2}x_{it}^2+\frac{\nu}{2}(p_{it}-p_{i,t-1})^2\Big],\qquad x_{it}=p_{it}-p_{it}^*. \tag{B.1}$$

The problem is linear-quadratic, so certainty equivalence holds and the first-order condition is

$$x_{it}+\nu\Big[(p_{it}-p_{i,t-1})-\beta\,\mathbb{E}_t(p_{i,t+1}-p_{it})\Big]=0. \tag{B.2}$$

Lemma B.1 (Rotemberg policy). *Let $\alpha_{\mathrm{R}}\in(0,1)$ be the smaller root of*

$$\beta a^2-(1+\beta+\frac{1}{\nu})a+1=0. \tag{B.3}$$

Then $1/\nu=\kappa(\alpha_{\mathrm{R}})$, with $\kappa(\alpha)\equiv(1-\alpha)(1-\beta\alpha)/\alpha$ as in Lemma 1, and the bounded solution of (B.2) *is*

$$p_{it}=\alpha_{\mathrm{R}}\,p_{i,t-1}+(1-\alpha_{\mathrm{R}})(1-\beta\alpha_{\mathrm{R}})\,\mathbb{E}_t\sum_{s\geq 0}(\beta\alpha_{\mathrm{R}})^s p_{i,t+s}^*,\qquad \textit{or}\qquad \mathbf{p}_i=J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})\,\mathbf{p}_i^*, \tag{B.4}$$

with $J^{\mathrm{C}}_{P,p^*}(\alpha)=(1-\alpha)(1-\beta\alpha)\Upsilon(\alpha)\Theta(\alpha)$ the Calvo desired-price pass-through of (24).

Proof. Substituting $x_{it}=p_{it}-p_{it}^*$ into (B.2) and dividing by ν,

$$\beta\,\mathbb{E}_t p_{i,t+1}-(1+\beta+\frac{1}{\nu})p_{it}+p_{i,t-1}=-\frac{1}{\nu}\,p_{it}^*. \tag{B.5}$$

The characteristic polynomial $\beta a^2-(1+\beta+1/\nu)a+1$ equals 1 at $a=0$ and $-1/\nu<0$ at $a=1$, so it has a root $\alpha_{\mathrm{R}}\in(0,1)$; the product of the roots is $1/\beta$, so the other root is $1/(\beta\alpha_{\mathrm{R}})>1$. Evaluating (B.3) at α_{R} and dividing by α_{R} gives $1/\nu=(1-\alpha_{\mathrm{R}})(1-\beta\alpha_{\mathrm{R}})/\alpha_{\mathrm{R}}=\kappa(\alpha_{\mathrm{R}})$, and

[23] As with the menu cost, taking the adjustment cost proportional to frictionless revenue, $\frac{\nu}{2}C_{it}(p_{it}-p_{i,t-1})^2$, makes ν constant across firms and dates once the objective is divided by C_{it}.

also $1+\beta+1/\nu=\beta\alpha_{\mathrm{R}}+1/\alpha_{\mathrm{R}}$. With the latter, (B.5) factors as

$$\beta(\mathbb{E}_t p_{i,t+1}-\alpha_{\mathrm{R}}p_{it})-\frac{1}{\alpha_{\mathrm{R}}}(p_{it}-\alpha_{\mathrm{R}}p_{i,t-1})=-\frac{1}{\nu}\,p^*_{it}.$$

Let $u_{it}\equiv p_{it}-\alpha_{\mathrm{R}}p_{i,t-1}$. Then $u_{it}=\beta\alpha_{\mathrm{R}}\,\mathbb{E}_t u_{i,t+1}+\dfrac{\alpha_{\mathrm{R}}}{\nu}\,p^*_{it}$, whose unique bounded solution is

$$u_{it}=\frac{\alpha_{\mathrm{R}}}{\nu}\,\mathbb{E}_t\sum_{s\geq 0}(\beta\alpha_{\mathrm{R}})^s p^*_{i,t+s};$$

the explosive root is excluded by boundedness. Since $\alpha_{\mathrm{R}}/\nu=\alpha_{\mathrm{R}}\kappa(\alpha_{\mathrm{R}})=(1-\alpha_{\mathrm{R}})(1-\beta\alpha_{\mathrm{R}})$, this is (B.4). In sequence space, $\mathbb{E}_t\sum_s(\beta\alpha_{\mathrm{R}})^s p^*_{i,t+s}=[\Theta(\alpha_{\mathrm{R}})\mathbf{p}^*_i]_t$, and inverting $p_{it}-\alpha_{\mathrm{R}}p_{i,t-1}=(1-\alpha_{\mathrm{R}})(1-\beta\alpha_{\mathrm{R}})[\Theta(\alpha_{\mathrm{R}})\mathbf{p}^*_i]_t$ with $\Upsilon(\alpha_{\mathrm{R}})=(I-\alpha_{\mathrm{R}}L)^{-1}$ gives

$$\mathbf{p}_i=(1-\alpha_{\mathrm{R}})(1-\beta\alpha_{\mathrm{R}})\,\Upsilon(\alpha_{\mathrm{R}})\Theta(\alpha_{\mathrm{R}})\,\mathbf{p}^*_i=J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})\,\mathbf{p}^*_i,$$

the map (24) with α_{R} in place of α. □

The price of a Rotemberg firm is thus the Calvo expected price (23) with α_{R} in place of α, firm by firm and not only in expectation: the reset formula $\Theta(\alpha_{\mathrm{R}})$ and the survival weights $\Upsilon(\alpha_{\mathrm{R}})$ arise here from the two roots of the Euler equation, not from an adjustment lottery.

Proposition B.1 (Rotemberg). *In the (R) economy,*

(*i*) *The desired-price pass-through Jacobian is Calvo:* $J^{\mathrm{R}}_{P,p^*}=J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})$*, and hence* $J^{\mathrm{R}}_{\pi,mc^r}=J^{\mathrm{C}}_{\pi,mc^r}(\alpha_{\mathrm{R}})=\omega\,\kappa(\alpha_{\mathrm{R}})\,\Theta_\beta=\dfrac{\omega}{\nu}\,\Theta_\beta$*, the New Keynesian Phillips curve* $\pi_t=\dfrac{\omega}{\nu}\,mc^r_t+\beta\mathbb{E}_t\pi_{t+1}$*;*

(*ii*) *Proposition 1 holds for any distribution of* ϵ_{it}*, and the IRF statistics are exact:* $b^{\mathrm{R}}_h=\omega(1-\alpha_{\mathrm{R}}^{h+1})$*;* $\phi^{\mathrm{IRF,R}}=\alpha_{\mathrm{R}}$*; and* $J^{\mathrm{IRF,R}}_{\pi,mc^r}=\kappa^{\mathrm{IRF,R}}\Theta_\beta=J^{\mathrm{R}}_{\pi,mc^r}$*;*

(*iii*) *The frequency of price changes is one, so* $1-\mathrm{freq}=0$ *while* $\phi^{\mathrm{IRF,R}}=\alpha_{\mathrm{R}}\in(0,1)$*: the frequency carries no information about the Phillips curve, and the Calvo slope built from it,* $\omega\kappa(1-\mathrm{freq})$*, is unbounded.*

Proof. (i) Integrating (B.4) across firms, the idiosyncratic components of $\mathbf{p}^*_i$ wash out as in Section 3.3.1, so $\mathbf{P}=J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})[\omega\mathbf{mc}+(1-\omega)\mathbf{P}]$ and $J^{\mathrm{R}}_{P,p^*}=J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})$. Lemma 1(ii) then gives $J^{\mathrm{R}}_{\pi,mc^r}=\omega\kappa(\alpha_{\mathrm{R}})\Theta_\beta$, and $\kappa(\alpha_{\mathrm{R}})=1/\nu$ by Lemma B.1.

(ii) Fix a firm, its predetermined state S and the other draws $\mathcal{D}$. By (B.4) the price path is affine in the path of desired prices, and under the random walk a shift δ in ϵ_{it}

raises $p^*_{i,t+s}$ by $\omega\delta$ for every $s \geq 0$ and nothing else; hence the cumulative price change $f_h(\epsilon_{it}; S, \mathcal{D})$ is affine in ϵ_{it} with slope $\omega[J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})\mathbf{1}]_h$, which does not depend on $(S, \mathcal{D})$. Case (ii) of Lemma B.4 therefore applies for any distribution of ϵ_{it}, and with Lemma B.3, $b^{\mathrm{R}}_h = \omega[J^{\mathrm{C}}_{P,p^*}(\alpha_{\mathrm{R}})\mathbf{1}]_h = \omega(1-\alpha_{\mathrm{R}}^{h+1})$ by Lemma 1(i). Proposition 2 then applies verbatim with $\alpha = \alpha_{\mathrm{R}}$: $\phi^{\mathrm{IRF,R}} = \alpha_{\mathrm{R}}$ and $J^{\mathrm{IRF,R}}_{\pi,mc^r} = \kappa^{\mathrm{IRF,R}}\Theta_\beta = \omega\kappa(\alpha_{\mathrm{R}})\Theta_\beta = J^{\mathrm{R}}_{\pi,mc^r}$.

(iii) Under a convex cost the optimal price moves whenever the desired price does, so every price changes every period and freq $= 1$. Since $\kappa(\alpha) \to \infty$ as $\alpha \to 0$, the Calvo slope evaluated at $1 - \text{freq} = 0$ is unbounded. □

The result reads like Proposition 3 with α_{R} in place of η_λ, for the opposite reason. In (GL) the shock leaves the expected price gap at a geometric rate because reconsideration is a lottery that does not depend on the gap. In (R) it leaves the gap at a geometric rate because every firm closes a constant fraction $1-\alpha_{\mathrm{R}}$ of its distance to the forward-looking target each period; there is no selection, and $\phi^{\mathrm{IRF,R}}$ is a pure measure of the speed of adjustment. In both models the statistic reads the decay of the shock in prices, which is what the Phillips curve depends on, while the frequency of price changes is uninformative: below the relevant hazard in (GL), and equal to one in (R).

B.3. Gertler–Leahy

B.3.1. Selection

Remember that the relevant gap is the one to the reset price, y, on which the band of Appendix A.2 is defined. For a firm hit at t, this gap is:

$$y_{it} = p_{i,t-1} - p^{\text{new}}_{it} = \Big[(p_{i,t-1} - \omega z_{i,t-1}) - (p^{\text{new}}_{it} - \omega z_{it})\Big] - \omega\epsilon^z_{it}, \tag{B.6}$$

where, by (27), $p^{\text{new}}_{it} - \omega z_{it}$ does not depend on the current draw. Since z is frozen between hits, the bracket is the gap the firm had after its last hit, in $[-\bar{x}, \bar{x}]$, plus the drift of $p^{\text{new}} - \omega z$ since then, bounded by m under (A2) for every firm whose last hit is within $k(m)$ periods, a set of mass at least $1 - cm$; so it is at most $\bar{x} + m$ in absolute value for those firms, and the complement contributes a second-order error to Lemma B.2, which therefore holds to the order of the approximation. Since $\omega\epsilon^z_{it}$ is uniform on a support of half-width $\varsigma/2 > 2\bar{x} + m$ by (A3), y_{it} is uniform on an interval of width ς that contains $[-\bar{x}, \bar{x}]$.

LEMMA B.2 (Selection). *Conditional on a hit at t, on the aggregate and shifter paths, and on the firm's history: (i) a firm without a free draw adjusts with probability $1 - 2\bar{x}/\varsigma$, and one with a free draw with certainty; (ii) $\mathbb{E}[p_{it} \mid \textit{hit}] = \mathbb{E}[p_{it}^{\text{new}} \mid \textit{hit}]$, the expectation being over the turbulence draw.*

PROOF. (i) A firm without a free draw keeps its price if $y_{it} \in [-\bar{x}, \bar{x}]$, an event of probability $2\bar{x}/\varsigma$ under the uniform distribution; a free draw makes adjustment costless. (ii) If the firm adjusts, freely or because its gap is outside the band, $p_{it} - p_{it}^{\text{new}} = 0$; if not, $p_{it} - p_{it}^{\text{new}} = y_{it}$ with $|y_{it}| \leq \bar{x}$. Hence $\mathbb{E}[p_{it} - p_{it}^{\text{new}} \mid \text{hit}] = (1-\lambda)\,\mathbb{E}[y_{it}\mathbf{1}\{|y_{it}| \leq \bar{x}\}] = (1-\lambda)\,\varsigma^{-1}\int_{-\bar{x}}^{\bar{x}} y\,dy = 0$. □

Uniformity makes the conditional density of non-adjusters flat, and symmetry of the band puts the reset price at its midpoint; both are needed. With a non-uniform shock the truncated mass would tilt toward one edge and the conditional mean would miss p_{it}^{new}.

B.3.2. The Calvo Law of Expected Prices and Proof of Proposition 3

The selection property (29) is Lemma B.2(ii). *The Calvo law* (30). Take expectations over the whole adjustment process (hits, free draws and turbulence draws), conditional on the paths of mc, P and ϑ_i. Partition firms at t into free adjusters (mass λ), whose price is the reset price; firms without a free draw that are hit (mass $(1-\lambda)(1-\eta)$), whose expected price is the expected reset price by Lemma B.2(ii); and the rest (mass η_λ), which keep their price, $p_{it} = p_{i,t-1}$. The expectation of the reset price over the turbulence history is $\bar{p}_{it}^{\text{new}}$, equation (28), because turbulence draws are mean zero, and the event "neither hit nor free" is independent of the firm's history, so the expected price of the third group is $\bar{p}_{i,t-1}$. Hence

$$\bar{p}_{it} = \eta_\lambda\,\bar{p}_{i,t-1} + (1-\eta_\lambda)\,\bar{p}_{it}^{\text{new}}, \tag{B.7}$$

which is (30).

Proposition 3, macro. By the law of large numbers across firms, $P_t = \int p_{it}\,di = \int \bar{p}_{it}\,di$, and since $\int z_{it}\,di = \int \vartheta_{it}\,di = 0$, integrating (B.7) across firms gives

$$P_t = \eta_\lambda\,P_{t-1} + (1-\eta_\lambda)(1-\eta_\lambda\beta)\sum_{s\geq 0}(\eta_\lambda\beta)^s\,[\omega\,mc_{t+s} + (1-\omega)P_{t+s}]. \tag{B.8}$$

In sequence form $\mathbf{P} = (1-\eta_\lambda)\Upsilon(\eta_\lambda)(1-\eta_\lambda\beta)\Theta(\eta_\lambda)[\omega\mathbf{mc} + (1-\omega)\mathbf{P}] = J^{\mathrm{C}}_{P,p^*}(\eta_\lambda)[\omega\mathbf{mc} + (1-\omega)\mathbf{P}]$, so the desired-price pass-through is $J^{\mathrm{GL}}_{P,p^*} = J^{\mathrm{C}}_{P,p^*}(\eta_\lambda)$ and $J^{\mathrm{GL}}_{P,mc} = \omega J^{\mathrm{C}}_{P,p^*}(\eta_\lambda)$. By Lemma 1(ii), $J^{\mathrm{GL}}_{\pi,mc^r} = \omega\kappa(\eta_\lambda)\Theta_\beta$, i.e. $\pi_t = \omega\kappa(\eta_\lambda)mc^r_t + \beta\pi_{t+1}$ with $\kappa(\eta_\lambda) = (1-\eta_\lambda)(1-\eta_\lambda\beta)/\eta_\lambda$.

Proposition 3, micro. By Lemma B.2, which holds for every realization of the shifter path under (A2)–(A3), the conditional expectation of $p_{i,t+h}$ given the shifter path is obtained by iterating (B.7) and is linear in that path. With a random walk, $\partial\bar{p}^{\mathrm{new}}_{i,t+j}/\partial\epsilon_{it} = \omega$ for $j \geq 0$ by (28), so

$$\frac{\partial\bar{p}_{i,t+h}}{\partial\epsilon_{it}} = \omega(1-\eta_\lambda)\sum_{j=0}^{h}\eta_\lambda^{h-j} = \omega(1-\eta_\lambda^{h+1}) = \omega[J^{\mathrm{C}}_{P,p^*}(\eta_\lambda)\mathbf{1}]_h, \tag{B.9}$$

and by linearity the regression coefficient equals this derivative, $b^{\mathrm{GL}}_h = \omega(1-\eta_\lambda^{h+1})$. Then $\phi^{\mathrm{IRF,GL}} = b_1/b_0 - 1 = (1-\eta_\lambda^2)/(1-\eta_\lambda) - 1 = \eta_\lambda$.

Frequency. A firm changes its price if and only if it draws a free adjustment, or is hit without one and adjusts: freq $= \lambda + (1-\lambda)(1-\eta)(1-2\bar{x}/\varsigma)$ by Lemma B.2(i), and $1-\text{freq} = \eta_\lambda + (1-\lambda)(1-\eta)\,2\bar{x}/\varsigma > \eta_\lambda$. □

B.4. Menu Cost

Every object in this appendix is computed in model (MC); to lighten notation we write ϕ^{IRF}, κ^{IRF} and b_h for $\phi^{\mathrm{IRF,MC}}$, $\kappa^{\mathrm{IRF,MC}}$ and b^{MC}_h.

B.4.1. The Optimal Policy and the Stationary Distribution

This subsection derives the Ss form of the policy stated in Section 2.1, for the normalized problem of Appendix A.1. In gap units the firm's state at the start of a period is its pre-adjustment gap $y = x_- - v$: the gap carried from the previous period, x_-, once the desired price has moved by the innovation $v = \omega(\epsilon^z_{it} + \epsilon_{it})$ and before any adjustment, where v is i.i.d. with a density f that is symmetric and single-peaked; the menu cost draw is $\xi_{it} \in \{0, \xi\}$ with $\Pr(\xi_{it} = 0) = \lambda$, and ξ is the normalized menu cost of Appendix A.1. Let $V(y)$ be the expected discounted loss of a firm with pre-adjustment gap y, before the

menu-cost draw, and

$$W(x) \equiv \frac{1}{2}x^2 + \beta\, \mathbb{E}_{v'}[V(x - v')] \tag{B.10}$$

the value of ending the period with gap x. The firm's problem is

$$V(y) = \lambda \min_x W(x) + (1-\lambda) \min\left\{W(y),\ \xi + \min_x W(x)\right\} \tag{B.11}$$

: with a free draw the firm chooses the best gap; otherwise it keeps y or pays ξ and chooses the best gap. Four properties follow.

(*i*) *Reset gap.* The right-hand side of (B.11) maps even functions into even functions when f is symmetric, since the static loss is even and $\mathbb{E}_{v'}V(x - v') = \mathbb{E}_{v'}V(-x - v')$ for even V; so V and W are even and the set of minimizers of W is symmetric. Auclert et al. (2024) show that for single-peaked f the minimizer is unique, hence $x^* = 0$: the reset gap is zero, and a firm that adjusts sets its price at the desired price.

(*ii*) *Band.* A firm without a free draw adjusts if and only if $W(y) - W(0) > \xi$. Since W is even and increasing in $|y|$ (single-peakedness of f propagates through the expectation), the adjustment region is $|y| > \bar{x}$, with the half-width defined by value matching,

$$W(\bar{x}) - W(0) = \xi, \tag{B.12}$$

so the inaction region is the symmetric band $[-\bar{x}, \bar{x}]$; in discrete time there is no smooth-pasting condition. Value matching (B.12) has no closed form in general, since $W(y) - W(0)$ includes the option of adjusting in later periods; but $W(y) - W(0) \approx \frac{1}{2}cy^2$ near zero, so a second-order menu cost gives a first-order band, $\bar{x} = O(\sqrt{\xi})$, which is what makes the gap first order and the remainder of (4) third order in the derivation of the objective (5) (Appendix A.1).

(*iii*) *Stationary distribution and frequency.* The post-adjustment gap is 0 for a firm that adjusts and y for one that does not, and the next pre-adjustment gap is the post-adjustment gap minus the next innovation. Writing μ for the stationary law of the post-adjustment gap and g for the stationary density of the pre-adjustment gap,

(B.13)

$$g(y) = \int f(y-x)\,\mu(dx), \qquad \mu = \text{freq}\,\delta_0 + (1-\lambda)\,g\,\mathbf{1}_{[-\bar{x},\bar{x}]}, \qquad \text{freq} = 1-(1-\lambda)\int_{-\bar{x}}^{\bar{x}} g(y)\,dy$$

: μ has an atom at the reset point and a density on the band, and g is a density (a convolution with f), as smooth as f. Written out, since μ has mass freq at zero and density $(1-\lambda)g$ on the band, the first equation in (B.13) is the integral equation

$$g(y) = \text{freq}\, f(y) + (1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(y-x)\, g(x)\, dx. \tag{B.14}$$

Moreover, the atom of μ at zero is the reason the cross-sectional law μ_S of Appendix B.5 has an atom, and the smoothness of g is what makes the aggregate response differentiable there. Existence and uniqueness of (μ, g) follow from the reinjection at zero, which makes the transition kernel a mixture with a fixed component, as Auclert et al. (2024) show.

(*iv*) *Smooth hazard.* With a menu cost drawn from a continuous distribution H, the firm adjusts if and only if $\xi_{it} < W(y) - W(0)$, so the adjustment probability given the gap is

$$\Lambda(y) = H(W(y) - W(0)), \tag{B.15}$$

even, increasing in $|y|$, and as smooth as H and W; the CalvoPlus case is $H = \lambda\delta_0 + (1-\lambda)\delta_\xi$, for which $\Lambda(y) = \lambda + (1-\lambda)\mathbf{1}\{|y| > \bar{x}\}$, and (B.13) holds with $\int \Lambda\, dg$ in place of the band integral.

B.4.2. Impact Entry of the Pass-Through Matrix

This subsection computes the first entry of the desired-price pass-through matrix, $[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0$: the date-$t$ response of the price level to a permanent unit shift of every firm's desired price from date t on, with the price level held fixed as in (11), evaluated at the stationary cross-section. Let $y \equiv x_{i,t-1} - v_{it}$ be the pre-adjustment gap, with stationary density g (equation (B.13)), and raise every desired price by δ from date t on. With the price level fixed, the shift leaves the normalized problem of Appendix B.4.1 unchanged: the reset gap remains $x^* = 0$, so an adjuster sets its price at the shifted desired price, and the band $[-\bar{x}, \bar{x}]$ is the same. Its only effect is to lower every pre-adjustment gap by δ, from y to $y - \delta$. A firm adjusts if it draws a free adjustment (probability λ) or if $|y - \delta| > \bar{x}$, and its price change is then $-(y - \delta)$; otherwise its price change is zero. The average price change at date t is

therefore

$$\Delta\bar{p}_t(\delta) = -\lambda \int (y-\delta)\, g(y)\, dy - (1-\lambda) \int_{|y-\delta|>\bar{x}} (y-\delta)\, g(y)\, dy, \tag{B.16}$$

and $[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0$ is its derivative at $\delta = 0$, since the pre-shift price level does not depend on δ. Differentiating (B.16) (Leibniz's rule, g being continuous) gives two terms. On the adjustment region the price change moves one for one with δ, which contributes the mass of adjusters, $\lambda + (1-\lambda)\int_{|y|>\bar{x}} g(y)\, dy = \text{freq}$. At the edges of the band the adjustment decision itself changes: firms with $y \in (\bar{x}, \bar{x}+\delta]$, of mass $(1-\lambda)g(\bar{x})\delta$, are pulled inside the band and forgo a price change of $-\bar{x}$, and firms with $y \in (-\bar{x}, -\bar{x}+\delta]$, of mass $(1-\lambda)g(-\bar{x})\delta$, are pushed outside and change their price by $+\bar{x}$; each group contributes $+\bar{x}$ per unit of mass. Hence

$$[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0 = \text{freq} + (1-\lambda)\,\bar{x}\,[g(\bar{x}) + g(-\bar{x})] = \text{freq} + 2(1-\lambda)\, g(\bar{x})\, \bar{x}, \tag{B.17}$$

using the symmetry of g: the intensive margin (adjusters move one for one with the desired price) plus the extensive margin (the shift sweeps a mass $g(\bar{x})$ of firms across each edge of the band, and each of them moves its price by $\bar{x}$). This is the index of macroeconomic flexibility of Caballero and Engel (2007). With a continuous menu-cost distribution and adjustment probability $\Lambda(y)$, equation (B.15), averaging over the menu-cost draw gives an expected price change of $-y\Lambda(y)$ for a firm with pre-adjustment gap y, and $-(y-\delta)\Lambda(y-\delta)$ after the shift, so that $\Delta\bar{p}_t(\delta) = -\int (y-\delta)\Lambda(y-\delta)g(y)\, dy$ and, differentiating under the integral sign,

$$[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0 = \int [\Lambda(y) + y\Lambda'(y)]g(y)\, dy = \text{freq} + \int y\, \Lambda'(y)\, g(y)\, dy, \tag{B.18}$$

which nests (B.17) when Λ' reduces to point masses of $+(1-\lambda)$ at $y = \bar{x}$ and $-(1-\lambda)$ at $y = -\bar{x}$. Proposition 1 gives $b_0 = \omega[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0$, equation (31).

Horizon-1 entry. The same argument carried one period further gives the second entry, $[J^{\text{MC}}_{P,p^*}\mathbf{1}]_1$, in closed form. Write $v = \omega(\epsilon^z_{it} + \epsilon_{it})$ for the desired-price innovation, with density f: symmetric and single-peaked as in Appendix B.4.1, but not Gaussian, since only ϵ_{it} is (the convolution of the symmetric single-peaked law of ϵ^z_{it} with a Gaussian is symmetric

and single-peaked). Only the symmetry and continuity of f are used below. Let

$$m(x) \equiv -\mathbb{E}_v[(x-v)\,\Lambda(x-v)] \tag{B.19}$$

be the expected price change next period of a firm that carries the post-adjustment gap x into the period: it adjusts with probability $\Lambda(x-v)$ and then closes the gap $x-v$. In the notation of the footnote to (33), $m(x) = E^1(x) - x$, so the one-period virtual survival values are $\Phi_1^e = E^1(\bar{x})/\bar{x} = 1 + m(\bar{x})/\bar{x}$ and $\Phi_1^i = E^{1\prime}(0) = 1 + m'(0)$, which is how m connects to the objects of Section 3.3.3. Since f is symmetric and Λ even, m is odd, $m(0) = 0$, and $m(\bar{x}) < 0$: a firm at the upper edge of the band expects to lower its price. Its derivative is

$$-m'(x) = \Lambda_1(x) + (1-\lambda)\,\bar{x}\,[f(x-\bar{x}) + f(x+\bar{x})], \qquad \Lambda_1(x) \equiv \mathbb{E}_v[\Lambda(x-v)], \tag{B.20}$$

the firm-level impact response to a displacement of its gap: its probability of adjusting next period, $\Lambda_1(x)$, plus the extensive margin generated by its own next-period gap density $f(\cdot - x)$ at the edges of the band. Averaging (B.20) over the stationary post-adjustment law μ of (B.13) and using $g(y) = \int f(y-x)\,\mu(dx)$ recovers (B.17): $\int(-m')\,d\mu = \text{freq} + (1-\lambda)\bar{x}[g(\bar{x}) + g(-\bar{x})]$.

Consider the shifted cross-section of date t. The post-adjustment law under the shift, μ_δ, has an atom at zero of mass $\text{freq}(\delta) = \lambda + (1-\lambda)\Pr(|y-\delta| > \bar{x})$ and, on the band, the density $(1-\lambda)g(x+\delta)$ of the non-adjusters, whose gap has been displaced by $-\delta$; by the symmetry of g, $\text{freq}'(0) = (1-\lambda)[g(-\bar{x}) - g(\bar{x})] = 0$, so to first order the shift only reshapes the band density, $\partial_\delta \mu_\delta = (1-\lambda)g'(x)\,dx$ on $[-\bar{x}, \bar{x}]$. The desired price is at its shifted level from t on and the problem is stationary again, so the average price change at $t+1$ is $\int m(x)\,\mu_\delta(dx)$, and its derivative at $\delta = 0$ is the increment of the pass-through between the two dates:

$$\begin{aligned}[J^{\text{MC}}_{P,p^*}\mathbf{1}]_1 - [J^{\text{MC}}_{P,p^*}\mathbf{1}]_0 &= (1-\lambda)\int_{-\bar{x}}^{\bar{x}} m(x)\,g'(x)\,dx \\ &= 2(1-\lambda)\,g(\bar{x})\,m(\bar{x}) + (1-\lambda)\int_{-\bar{x}}^{\bar{x}} [-m'(x)]\,g(x)\,dx,\end{aligned} \tag{B.21}$$

by integration by parts, m being odd and g even. Substitute (B.20) and use, with $F_v(\bar{x}) \equiv$

$\Pr(|v| \leq \bar{x})$ and the convolution identity for g,

$$(1-\lambda)\int_{-\bar{x}}^{\bar{x}} \Lambda_1(x)\, g(x)\, dx = \Pr(\text{keep at } t,\ \text{adjust at } t+1) = \text{freq}\,(1-\lambda)\, F_v(\bar{x}),$$
$$(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x \mp \bar{x})\, g(x)\, dx = g(\pm\bar{x}) - \text{freq}\, f(\pm\bar{x}),$$

so that

$$[J^{\text{MC}}_{P,p^*}\mathbf{1}]_1 = [J^{\text{MC}}_{P,p^*}\mathbf{1}]_0 + \underbrace{\text{freq}\,(1-\lambda)\, F_v(\bar{x})}_{\text{intensive margin at } t+1} + \underbrace{2(1-\lambda)\,\bar{x}\,[g(\bar{x}) - \text{freq}\, f(\bar{x})]}_{\text{extensive margin at } t+1} - \underbrace{2(1-\lambda)\, g(\bar{x})\, |m(\bar{x})|}_{\text{unwinding of the date-}t\text{ extensive margin}}. \tag{B.22}$$

Only two groups of firms respond at $t+1$. Date-t adjusters sit at gap zero relative to the shifted desired price and behave as in the steady state, so they contribute nothing further. Date-t non-adjusters carry a gap displaced by $-\delta$ and produce at $t+1$ a fresh impact response: their adjustment probability, the intensive term, plus $\bar{x}$ times the mass they place at the edges of the band, the extensive term; their contribution to the density at $\pm\bar{x}$ is $g(\bar{x})$ net of the part $\text{freq} f(\bar{x})$ that comes from the atom at zero, which is why the extensive margin at $t+1$ is smaller than at t. Finally, the marginal firms of date t are not where they would have been: the firm pulled inside the band at $+\bar{x}$ now expects a price change $m(\bar{x}) < 0$ instead of zero, and the firm pushed out at $-\bar{x}$ was reset to zero instead of carrying the expected change $m(-\bar{x}) > 0$; both undo part of the date-t extensive response, which is why the survival function attached to the extensive margin in the two-margin representation (33) is virtual.

Equation (B.22) is the two-margin representation at horizon one. With the virtual survival functions of the footnote to (33), $\Phi^i_s = E^{s\prime}(0)$ and $\Phi^e_s = E^s(\bar{x})/\bar{x}$, and since $E^1(x) = x + m(x)$ (the expected gap next period is the current gap plus the expected price change, the innovation having mean zero),

$$\Phi^i_1 = 1 + m'(0) = (1-\lambda)F_v(\bar{x}) - 2(1-\lambda)\,\bar{x}\, f(\bar{x}), \qquad \Phi^e_1 = 1 + \frac{m(\bar{x})}{\bar{x}}, \tag{B.23}$$

and regrouping the terms of (B.22) with the help of (B.17),

$$[J^{\mathrm{MC}}_{P,p^*}\mathbf{1}]_1 = \text{freq}\,(1+\Phi^i_1) + 2(1-\lambda)\,g(\bar{x})\,\bar{x}\,(1+\Phi^e_1), \tag{B.24}$$

which is Lemma 2 applied to (33) at $h = 1$, i.e. (34) with $b_1/\omega = (b_0/\omega)(1+\bar{\Phi}_1)$. Two remarks. First, Φ^i_1 is not the plain one-period survival probability $(1-\lambda)F_v(\bar{x})$ of a freshly reset price: it carries the firm-level extensive correction $-2(1-\lambda)\bar{x}f(\bar{x})$, because the intensive component of the representation is the derivative of the expected gap at the reset point, and that derivative includes the mass of freshly reset firms that the next innovation places at the edge of the band. The two groupings, (B.22) and (B.24), distribute the same terms differently and sum to the same number. Second, in primitives the statistic of Proposition 4 is

$$\phi^{\mathrm{IRF}} = \frac{b_1}{b_0} - 1 = \bar{\Phi}_1 = \frac{\text{freq}\,\Phi^i_1 + 2(1-\lambda)\,g(\bar{x})\,\bar{x}\,\Phi^e_1}{\text{freq} + 2(1-\lambda)\,g(\bar{x})\,\bar{x}}, \tag{B.25}$$

with Φ^i_1 and Φ^e_1 given by (B.23); the extensive margin typically decays much faster than the intensive one ($\Phi^e_1 < \Phi^i_1$), since a firm that starts at the threshold adjusts with probability at least one half at the next draw: any innovation of the wrong sign pushes it out of the band. The same recursion holds at every horizon, $[J^{\mathrm{MC}}_{P,p^*}\mathbf{1}]_h - [J^{\mathrm{MC}}_{P,p^*}\mathbf{1}]_{h-1} = \text{freq}\,\Phi^i_h + 2(1-\lambda)g(\bar{x})\bar{x}\,\Phi^e_h$, but for $h \geq 2$ the functions E^h are iterates of the one-step map $x \mapsto x + m(x)$ composed with the adjustment policy and have no closed form in $(g, f, \bar{x})$ alone: horizon one is the last entry with an explicit expression, which is also why Section 3.3.3 extrapolates from (b_0, b_1) rather than computing further entries.

Comparison with $1-$freq: selection versus the hazard. Section 3.3.3 states that $\phi^{\mathrm{IRF}} = \bar{\Phi}_1$ falls short of $1 -$ freq, the survival probability of a price spell. The closed forms above show exactly when and why. Let $q(x) \equiv \Pr_v(|x-v| > \bar{x})$ be the probability that a firm carrying the post-adjustment gap x into the next period is pushed out of the band, so that $\Lambda_1(x) = \lambda + (1-\lambda)q(x)$. Since f is symmetric and decreasing on $\mathbb{R}_+$, q is even and increasing in $|x|$: the threshold firm is the most exposed to exit, the reset firm the least. Averaging Λ_1 over the post-adjustment law μ of (B.13) gives the frequency, $\text{freq} = \lambda + (1-\lambda)\int q\,d\mu$, and

$F_v(\bar{x}) = 1 - q(0)$. By (B.25),

$$[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0\,[\phi^{\text{IRF}} - (1-\text{freq})] = \text{freq}\,[\Phi_1^i - (1-\text{freq})] + 2(1-\lambda)\,g(\bar{x})\,\bar{x}\,[\Phi_1^e - (1-\text{freq})], \tag{B.26}$$

so $\phi^{\text{IRF}} - (1 - \text{freq})$ has the sign of a weighted sum of the two component gaps, weighted by the intensive and extensive masses of (B.17).

The extensive component is always below $1-\text{freq}$. From (B.23), $1-\Phi_1^e = |m(\bar{x})|/\bar{x}$. Splitting the expectation in $m(\bar{x})$ according to whether the innovation pushes the threshold firm out of the band ($v < 0$ or $v > 2\bar{x}$) or leaves it inside, and using the symmetry of f,

$$\begin{aligned} 1 - \Phi_1^e &= \lambda + (1-\lambda)\,q(\bar{x}) + \frac{1-\lambda}{\bar{x}} \int_0^{2\bar{x}} v\, f(v)\, dv, \\ \text{hence} \qquad \Phi_1^e - (1 - \text{freq}) &= -(1-\lambda)\Big[q(\bar{x}) - \int q\, d\mu + \frac{1}{\bar{x}} \int_0^{2\bar{x}} v\, f(v)\, dv\Big] < 0. \end{aligned} \tag{B.27}$$

Both terms in the bracket are selection. The threshold firm exits with the highest probability in the cross-section, $q(\bar{x}) \geq \int q\, d\mu$, whereas the frequency averages the exit probability over all post-adjustment gaps; and even when it stays inside the band, its gap mean-reverts, by the fraction $\frac{1}{\bar{x}} \int_0^{2\bar{x}} v f(v)\, dv$ of $\bar{x}$, because the innovations that keep it inside are the positive ones. Since $q(\bar{x}) \geq \Pr(v < 0) = \dfrac{1}{2}$, also $\Phi_1^e < (1-\lambda)/2$: a marginal gap at the threshold loses more than half of itself in one period, whatever the innovation density.

The intensive component can be on either side. From (B.23), $\Phi_1^i = (1-\lambda)[1 - q(0)] - 2(1-\lambda)\bar{x} f(\bar{x})$, so

$$\Phi_1^i - (1 - \text{freq}) = (1-\lambda)\Big[\underbrace{\int q\, d\mu - q(0)}_{\text{increasing hazard}} - \underbrace{2\bar{x}\, f(\bar{x})}_{\text{reset-point selection}}\Big]. \tag{B.28}$$

The first term is positive: a freshly reset price, at gap zero, exits less often than the average price, which is the increasing-hazard property of Ss models, and by itself it pushes the survival of the intensive margin, and hence ϕ^{IRF}, above $1 - \text{freq}$, exactly as an increasing hazard does in a time-dependent model (in a two-period Taylor model, $\Phi_1 = 1$ while $1 - \text{freq} = \dfrac{1}{2}$). The second term is selection again, now generated by the freshly reset firms themselves: the mass $2\bar{x} f(\bar{x})$ of them that the next innovation places at the edges of the band, each moving by $\bar{x}$. Which term dominates depends on the width of the band

relative to the innovation: for narrow and medium bands ($\bar{x}$ up to about twice the standard deviation of v with Gaussian innovations) selection dominates and $\Phi_1^i < 1-\text{freq}$; for wide bands the hazard term dominates, $\Phi_1^i > 1-\text{freq}$, and the sign of (B.26) is decided by the weight of the extensive component.

Sufficient conditions. Write $H \equiv (1-\lambda)[\int q\,d\mu - q(0)] = (1-\lambda)^2 \int_{-\bar{x}}^{\bar{x}} [q(x)-q(0)]g(x)\,dx$ for the hazard term and $w \equiv 2(1-\lambda)g(\bar{x})\bar{x}$ for the extensive mass. Then $H \le w$: for $|x| \le \bar{x}$, $q(x)-q(0) = \int_0^{|x|}[f(\bar{x}-s)-f(\bar{x}+s)]\,ds \le \int_0^{|x|} f(\bar{x}-s)\,ds \le \bar{x}\,f(\bar{x}-|x|)$, because f is decreasing on $\mathbb{R}_+$, so that

$$
\begin{aligned}
H &\le (1-\lambda)^2\,\bar{x}\int_{-\bar{x}}^{\bar{x}} f(\bar{x}-|x|)\,g(x)\,dx = 2(1-\lambda)^2\,\bar{x}\int_0^{\bar{x}} f(\bar{x}-x)\,g(x)\,dx \\
&\le 2(1-\lambda)^2\,\bar{x}\int_{-\bar{x}}^{\bar{x}} f(\bar{x}-x)\,g(x)\,dx \le 2(1-\lambda)\,\bar{x}\,g(\bar{x}) = w,
\end{aligned}
$$

the last step by the convolution identity $g(\bar{x}) = \text{freq} f(\bar{x}) + (1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(\bar{x}-x)g(x)\,dx$. Substituting (B.27) and (B.28) into (B.26), with $2(1-\lambda)\bar{x}f(\bar{x}) = w\,f(\bar{x})/g(\bar{x})$,

$$
\begin{aligned}
[J^{\text{MC}}_{P,p^*}\mathbf{1}]_0\,[\phi^{\text{IRF}} - (1-\text{freq})] &= \text{freq}\,H - \text{freq}\,w\,\frac{f(\bar{x})}{g(\bar{x})} - w\,[1-\Phi_1^e-\text{freq}] \\
&\le w\Big\{\text{freq}\Big[2-\frac{f(\bar{x})}{g(\bar{x})}\Big] - (1-\Phi_1^e)\Big\}.
\end{aligned} \tag{B.29}
$$

Therefore $\phi^{\text{IRF}} < 1-\text{freq}$ whenever either

(a) reset-point selection dominates the hazard, $2\bar{x}f(\bar{x}) \ge \int q\,d\mu - q(0)$: then both component gaps are nonpositive by (B.27)–(B.28), and $\phi^{\text{IRF}} < 1-\text{freq}$ as a weighted average; or

(b) $\text{freq}\,[2 - f(\bar{x})/g(\bar{x})] \le 1-\Phi_1^e$, by (B.29); since $1-\Phi_1^e > (1+\lambda)/2$ by (B.27) and $f(\bar{x})/g(\bar{x}) \ge 0$, this holds in particular whenever $\text{freq} \le (1+\lambda)/4$.

Condition (a) covers narrow and medium bands, where the frequency is high, and condition (b) covers wide bands, where it is low. On a grid of configurations with Gaussian and Laplace innovations, $\bar{x}$ between one quarter and five standard deviations of v and λ up to 0.3, the inequality $\phi^{\text{IRF}} < 1-\text{freq}$ holds throughout, with a gap $(1-\text{freq}) - \phi^{\text{IRF}}$ between 0.04 and 0.48 that closes only toward the Calvo limit (wide band and large λ), and condition (b) with the exact value of $1-\Phi_1^e$ holds at every point of the grid.

In the language of Section 3.3.3: $(1-\text{freq}) - \phi^{\text{IRF}}$ is the threshold-selection term of

(B.27), weighted by the extensive mass, plus the reset-point selection term of (B.28) net of the increasing-hazard term, weighted by the frequency. It measures selection net of the duration dependence of the hazard; a model with an increasing hazard but no selection, such as a Taylor model, would place ϕ^{IRF} above $1 - \text{freq}$.

Consistency with the pass-through mixture (33)**.** Equation (B.24) was already identified above as Lemma 2 applied to (33) at $h = 1$; the same argument, run once for general h, recovers both closed forms of this subsection together. Lemma 2 gives $[J^{\text{TD}}_{P,p^*}(\Phi^k)\mathbf{1}]_h = \sum_{s=0}^{h} \Phi^k_s / \sum_{s\geq 0} \Phi^k_s$ for $k \in \{e, i\}$, so that

$$[J^{\text{MC}}_{P,p^*}\mathbf{1}]_h = \varrho_e \frac{\sum_{s=0}^{h} \Phi^e_s}{\sum_{s\geq 0} \Phi^e_s} + \varrho_i \frac{\sum_{s=0}^{h} \Phi^i_s}{\sum_{s\geq 0} \Phi^i_s}.$$

Since $E^0(x) = x$, $\Phi^e_0 = \Phi^i_0 = 1$, and the mixture weights $\varrho_e = 2(1-\lambda)g(\bar{x})\bar{x}\sum_s \Phi^e_s$, $\varrho_i = \text{freq}\sum_s \Phi^i_s$ cancel the two denominators at every horizon, leaving

$$[J^{\text{MC}}_{P,p^*}\mathbf{1}]_h = 2(1-\lambda)g(\bar{x})\bar{x}\sum_{s=0}^{h} \Phi^e_s + \text{freq}\sum_{s=0}^{h} \Phi^i_s.$$

At $h = 0$ this is (B.17), and at $h = 1$ it is (B.24). This illustrates that the derivation of this appendix and the mixture representation of Auclert et al. (2024) describe the same object rather than two objects that happen to coincide at b_0.

B.4.3. Proof of Proposition 4

Writing $e \equiv E\mathbf{1}$ and using $b_h = \omega[J^{\text{MC}}_{P,p^*}\mathbf{1}]_h$ (Proposition 1) together with Lemma 1(i),

$$\frac{b_h}{\omega} = 1 - (\alpha^\star)^{h+1} + e_h, \qquad \text{so} \qquad \frac{b_0}{\omega} = 1 - \alpha^\star + e_0, \quad \frac{b_1}{\omega} = 1 - (\alpha^\star)^2 + e_1. \tag{B.30}$$

The real rigidity cancels from the ratio b_1/b_0. Substituting into $\phi^{\text{IRF}} = b_1/b_0 - 1$ and subtracting $\alpha^\star$,

$$\phi^{\text{IRF}} - \alpha^\star = \frac{e_1 - (1+\alpha^\star)\, e_0}{1 - \alpha^\star + e_0}. \tag{B.31}$$

Under Assumption 1, $|e_0| \leq \bar{e}$ and $|e_1| \leq \bar{e}$ (since $|[E\mathbf{1}]_t| \leq \|E\|_H$ for $t \leq H$, with $H \geq 1$), and the denominator of (B.31) is at least $1 - \alpha^\star - \bar{e} > 0$, so

$$\left|\phi^{\text{IRF}} - \alpha^\star\right| \leq \frac{(2 + \alpha^\star)\,\bar{e}}{1 - \alpha^\star - \bar{e}}. \tag{B.32}$$

In particular, $b_0/\omega - (1 - \phi^{\text{IRF}}) = e_0 + (\phi^{\text{IRF}} - \alpha^\star) = O(\bar{e})$, which gives part (iii): $\hat{\omega} - \omega = \omega\,[b_0/\omega - (1 - \phi^{\text{IRF}})]/(1 - \phi^{\text{IRF}})$, and (B.32) with $|e_0| \leq \bar{e}$ bounds it as in (**??**). The interval $I_\alpha = [\min\{\alpha^\star, \phi^{\text{IRF}}\}, \max\{\alpha^\star, \phi^{\text{IRF}}\}]$ satisfies $I_\alpha \subset (0,1)$ as long as $\bar{\Phi}_1 \in (0,1)$, as in every calibration we consider, since $\phi^{\text{IRF}} = \bar{\Phi}_1$; on it, $M_H(I_\alpha) = \sup_{\alpha \in I_\alpha} \|\partial J^{\text{C}}_{P,p^*}(\alpha)/\partial\alpha\|_H$ is finite. Writing the difference of the two Calvo Jacobians as an integral and taking norms (the mean-value inequality),

$$\|J^{\text{C}}_{P,p^*}(\phi^{\text{IRF}}) - J^{\text{C}}_{P,p^*}(\alpha^\star)\|_H = \Big\| \int_{\alpha^\star}^{\phi^{\text{IRF}}} \frac{\partial J^{\text{C}}_{P,p^*}(\alpha)}{\partial\alpha}\, d\alpha \Big\|_H \leq M_H(I_\alpha)\left|\phi^{\text{IRF}} - \alpha^\star\right|.$$

Combining with (B.32) and the triangle inequality applied to $J^{\text{MC}}_{P,p^*} - J^{\text{C}}_{P,p^*}(\phi^{\text{IRF}}) = E + J^{\text{C}}_{P,p^*}(\alpha^\star) - J^{\text{C}}_{P,p^*}(\phi^{\text{IRF}})$,

$$\|J^{\text{MC}}_{P,p^*} - J^{\text{C}}_{P,p^*}(\phi^{\text{IRF}})\|_H \leq \|E\|_H + M_H(I_\alpha)\left|\phi^{\text{IRF}} - \alpha^\star\right| \leq \bar{e} + M_H(I_\alpha)\,\frac{(2 + \alpha^\star)\,\bar{e}}{1 - \alpha^\star - \bar{e}},$$

which is (38).

Parts (ii) and (iii). Write $J \equiv J^{\text{MC}}_{P,p^*}$ and $\tilde{J} \equiv J^{\text{C}}_{P,p^*}(\phi^{\text{IRF}})$, both truncated at the reporting horizon H. On $(H+1) \times (H+1)$ matrices the maximum-absolute-row-sum norm is the operator norm induced by the ℓ^∞ vector norm, hence submultiplicative, and $\|I - L\| \leq 2$. For the truncated Calvo matrix, $\|\tilde{J}\| = \max_{h \leq H}(1 - (\phi^{\text{IRF}})^{\,h+1}) < 1$ (its entries are positive with row sums $1 - (\phi^{\text{IRF}})^{\,h+1}$), so $I - \tilde{J}$ is invertible; invertibility of the truncated $I - J$ is assumed in the statement (it holds in all our computations, and whenever $\|J\| < 1$). Using $(I - A)^{-1}A = (I - A)^{-1} - I$ for both matrices and differencing,

$$(I - \tilde{J})^{-1}\tilde{J} - (I - J)^{-1}J = (I - \tilde{J})^{-1}\,(\tilde{J} - J)\,(I - J)^{-1}. \tag{B.33}$$

By (12)–(13), $J^{\mathrm{MC}}_{\pi,mc^r} - J^{\mathrm{C}}_{\pi,mc^r}(\phi^{\mathrm{IRF}}) = -\omega(I-L)[(I-\tilde{J})^{-1}\tilde{J} - (I-J)^{-1}J]$, so taking norms and using submultiplicativity,

$$\|J^{\mathrm{MC}}_{\pi,mc^r} - J^{\mathrm{C}}_{\pi,mc^r}(\phi^{\mathrm{IRF}})\| \le \omega\,\|I-L\|\,\|(I-\tilde{J})^{-1}\|\,\|\tilde{J}-J\|\,\|(I-J)^{-1}\| \le 2\,\omega\,\|(I-\tilde{J})^{-1}\|\,\|(I-J)^{-1}\|\,B_H,$$

where the last step uses that the truncated difference obeys $\|\tilde{J} - J\| \le \|J^{\mathrm{C}}_{P,p^*}(\phi^{\mathrm{IRF}}) - J^{\mathrm{MC}}_{P,p^*}\|_H \le B_H$: its row sums over $s \le H$ are bounded by the corresponding row sums over all s, and part (i) applies. This is (39).

B.5. Proof of Proposition 1

Expectations. Fix a date t and a firm i. Everything random about the firm's prices from t to $t+h$ is one of three objects: its predetermined state S (the gap $x_{i,t-1}$, the predetermined price $p_{i,t-1}$, and $\vartheta_{i,t-1}$ in the AR(1) case); its identifying innovation ϵ_{it}; and $\mathcal{D}$, all other random draws affecting its prices from t to $t+h$ (the innovations $\epsilon^z_{i,t+j}$ and $\epsilon_{i,t+j}$ for $j \ge 1$, the date-t innovation ϵ^z_{it}, and the menu-cost or lottery draws). Their laws are $S \sim \mu_S$, the cross-sectional law of the state at date t, stationary in its gap component (price levels have no stationary law, but $p_{i,t-1}$ enters every object below only additively and cancels from every cumulative price change); $\epsilon_{it} \sim \mu_\epsilon$, with density φ_σ in the Gaussian case; and $\mathcal{D} \sim \mu_{\mathcal{D}}$. The three are mutually independent: ϵ_{it} is independent of $(S, \mathcal{D})$ by the maintained assumption of (8) (i.i.d. over time, independent of the z process and of all adjustment draws), and $\mathcal{D}$ consists of date-t draws other than ϵ_{it} and of later innovations and draws, all independent of S. We denote expectations with a subscript naming the law it integrates over, the remaining arguments being held fixed: $\mathbb{E}_\epsilon$ integrates against μ_ϵ alone, $\mathbb{E}_{S,\mathcal{D}}$ against $\mu_S \otimes \mu_{\mathcal{D}}$, $\mathbb{E}_S$ against μ_S, and

$$\mathbb{E}_{S,\epsilon,\mathcal{D}} \equiv \mathbb{E}_{S,\mathcal{D}}\,\mathbb{E}_\epsilon = \mathbb{E}_\epsilon\,\mathbb{E}_{S,\mathcal{D}} \tag{B.34}$$

integrates against the product law $\mu_S \otimes \mu_\epsilon \otimes \mu_{\mathcal{D}}$. This last one is the *cross-sectional* expectation: the cross-section of firms at date t is the ergodic law of a single firm's $(S, \epsilon_{it}, \mathcal{D})$, so averaging over the distribution of firms means integrating against $\mu_S \otimes \mu_\epsilon \otimes \mu_{\mathcal{D}}$. Variances, covariances and probabilities carry the same subscripts.

Road map. The proof has two steps, stated as two lemmas. Conditional on $(S, \mathcal{D})$, the firm's cumulative price change is a deterministic function of its innovation, $p_{i,t+h} - p_{i,t-1} = f_h(\epsilon_{it}; S, \mathcal{D})$. Define the cross-sectional average derivative of the price path with respect to the firm's own shifter,

$$A_h \equiv \frac{d}{d\delta} \mathbb{E}_{S,\epsilon,\mathcal{D}}[f_h(\epsilon_{it} + \delta; S, \mathcal{D})]\Big|_{\delta=0}, \tag{B.35}$$

the derivative of the average cumulative price change when every firm's identifying innovation is raised by δ. Step 1 (Section B.5.1, Lemma B.3) shows $[J^m_{P,mc}\mathbf{1}]_h = A_h$ in every model, using only the structure of the firm's problem, the random-walk structure of the shifter, and no property of the marginal law of ϵ_{it}. Step 2 (Section B.5.2, Lemma B.4) shows $A_h = b^m_h$: the regression coefficient equals the average derivative. This is where the distribution of the identifying shock matters, and where Gaussianity enters for the menu cost model. Proposition 1 is the two lemmas combined.

B.5.1. Step 1: The Aggregate Response Is the Average Derivative

Fix a date t and a firm i, with S, ϵ_{it}, $\mathcal{D}$ and f_h as defined above. A shift of ϵ_{it} by δ is the experiment behind A_h in (B.35).

Lemma B.3 (Aggregation). *In each model of Section 2.1, with a random-walk shifter, let $R_h(\delta)$ be the mean price at $t+h$ when nominal marginal cost is shifted by δ from date t on, so that* every *firm's desired price is shifted by $\omega\delta$, with the price level and all other aggregates at their steady-state values and no feedback from prices to costs, as in* (11)*, starting from the stationary cross-section. Then, for every δ (in* (GL)*, for every δ within the slack of the small-shock bounds* (A2)–(A3)*),*

$$R_h(\delta) = \mathbb{E}_S[p_{i,t-1}] + \mathbb{E}_{S,\epsilon,\mathcal{D}}[f_h(\epsilon_{it} + \delta; S, \mathcal{D})], \tag{B.36}$$

and, whenever R_h is differentiable at 0 (equivalently, by (B.36)*, whenever A_h exists),*

$$[J^m_{P,mc}\mathbf{1}]_h = R_h'(0) = A_h. \tag{B.37}$$

Beyond the random-walk structure of ϑ and the independence of ϵ_{it} from later innovations and adjustment draws, which keep the continuation problem stationary, no property of the law of ϵ_{it} is used: its shape, moments and support, and its dependence on the predetermined state, play no role.

PROOF. Fix $(S, \epsilon_{it}, \mathcal{D})$ and a real δ, and compare two experiments, both started from the same state and driven by the same draws:

(E1) the firm's own innovation is $\epsilon_{it} + \delta$ instead of ϵ_{it}; nominal marginal cost stays at its steady-state value;

(E2) the firm's innovation is ϵ_{it}; nominal marginal cost rises permanently and unexpectedly by δ at date t, with the price level and all other aggregates held at their steady-state values as in (11).

Write $p^{(1)}_{i,t+h}(\delta)$ and $p^{(2)}_{i,t+h}(\delta)$ for the resulting prices at $t+h$; both are deterministic functions of $(S, \epsilon_{it}, \mathcal{D}, \delta)$.

(*a*) *Pathwise equivalence:* $p^{(1)}_{i,t+h}(\delta) = p^{(2)}_{i,t+h}(\delta)$ *for every* $(S, \epsilon_{it}, \mathcal{D})$ *and every* δ. By (5)—the normalized problem of Appendix A.1, in which the menu cost is proportional to frictionless revenue, so that a common cost shift leaves the normalized menu cost unchanged—the firm's objective and its adjustment technology depend on $(mc_\tau, z_{i\tau}, \vartheta_{i\tau}, P_\tau)$ only through the desired price $p^*_{i\tau}$, and on $p^*_{i\tau}$ only through the gap $x_{i\tau} = p_{i\tau} - p^*_{i\tau}$; its optimal policy at any date $\tau \geq t$ is therefore a function of the pre-adjustment gap at τ and of the conditional law of the future increments $\{p^*_{i,\tau+s} - p^*_{i\tau}\}_{s\geq 1}$. In (E1), $p^*_{i,t+s}$ rises by $\omega\delta$ for every $s \geq 0$ because ϑ is a random walk and enters (7) with weight ω, and the law of the future increments is unchanged because ϵ_{it} is independent of all later innovations and draws. In (E2), $p^*_{i,t+s}$ rises by the same $\omega\delta$ for every $s \geq 0$, since mc enters (7) with the same weight and the price level is held fixed, and the law of the future increments is unchanged because the shift is permanent and not revised. Hence in both experiments the pre-adjustment gap at t is $x_{i,t-1} - v_{it} - \omega\delta$, the desired price at $t+h$ is $p^*_{i,t+h} + \omega\delta$, and the continuation from t is solved by the stationary policies (bands, reset gap $x^* = 0$, hazard) applied to the same draws in $\mathcal{D}$. The realized price is therefore, in both experiments,

$$\text{(B.38)}\quad p_{i,t+h} = p^*_{i,t+h} + \omega\delta + \mathcal{X}_h(x_{i,t-1} - v_{it} - \omega\delta;\ v_{i,t+1}, \ldots, v_{i,t+h};\ \text{adjustment draws}),$$

where $v_{i\tau} = \omega(\epsilon^z_{i\tau} + \epsilon_{i\tau})$ is the date-τ innovation to the desired price (with $z_{i\tau} - z_{i,\tau-1}$ in place of $\epsilon^z_{i\tau}$ when z is not a random walk) and $\mathcal{X}_h$ is the gap h periods later under the stationary policies; when those policies condition on further state variables (z_{it} in (C) if z is not a random walk, ϑ_{it} in the AR(1) case), these enter $\mathcal{X}_h$ as additional arguments, none of which depends on δ. In (GL) the argument holds for $|\delta|$ within the slack of the small-shock bounds (A2)–(A3), under which the policies of Appendix A.2 remain the firm's

policies.

Intuitively, neither shock has an existence of its own in the firm's problem: both are the same shift $\omega\delta$ of $p^*_{it} = \omega(mc_t + z_{it} + \vartheta_{it}) + (1-\omega)P_t$, so the two pass-throughs are instances of a single object, the pass-through of the desired price, scaled by the common weight ω.

(*b*) *Aggregation.* The cross-section at date t is the law $\mu_S \otimes \mu_\epsilon \otimes \mu_{\mathcal{D}}$ of $(S, \epsilon_{it}, \mathcal{D})$ in both experiments: a common cost shift alters no idiosyncratic draw. Since the price level and aggregate output (P_t, Y_t) are held fixed, the shifted cost mc_t is common to all firms, and, given these, no firm's problem involves any other firm's price—with decreasing returns the firm's own output enters its own marginal cost, but that is a function of its own price and of the aggregates, and Appendix A.1 shows it is absorbed in the profit function $R^*_{it} h_\omega(x_{it})$, i.e. in the weight ω of the desired price and in the curvature—firms respond independently, and the mean price in (E2) is the integral of $p^{(2)}_{i,t+h}(\delta)$ against that law. Therefore

$$R_h(\delta) = \mathbb{E}_{S,\epsilon,\mathcal{D}}[p^{(2)}_{i,t+h}(\delta)] = \mathbb{E}_{S,\epsilon,\mathcal{D}}[p^{(1)}_{i,t+h}(\delta)] = \mathbb{E}_S[p_{i,t-1}] + \mathbb{E}_{S,\epsilon,\mathcal{D}}[f_h(\epsilon_{it} + \delta; S, \mathcal{D})],$$

the middle equality by (a) and the last by the definition of f_h, since $p^{(1)}_{i,t+h}(\delta) = p_{i,t-1} + f_h(\epsilon_{it} + \delta; S, \mathcal{D})$; this is (B.36). Subtracting $\mathbb{E}_S[p_{i,t-1}]$, which does not depend on δ, and differentiating at $\delta = 0$ gives $R'_h(0) = A_h$ whenever either side exists. Finally, $[J^m_{P,mc}\mathbf{1}]_h$ is by definition the derivative of the aggregate price at horizon h with respect to a permanent unit shift of nominal marginal cost at date 0, unanticipated and with aggregates fixed, starting from the stationary cross-section; relabelling date 0 as t is admissible because the cross-section is the same at every date. Hence $[J^m_{P,mc}\mathbf{1}]_h = R'_h(0) = A_h$, which is (B.37). □

When the derivative can be moved inside the expectation over $(S, \mathcal{D})$, equation (B.37) reads $[J^m_{P,mc}\mathbf{1}]_h = \mathbb{E}_{S,\mathcal{D}}[\frac{d}{d\delta}\mathbb{E}_\epsilon[f_h(\epsilon+\delta; S, \mathcal{D})]\big|_{\delta=0}]$. This is the aggregation logic of Caballero and Engel (2007): the cross-section over which the aggregate response averages is the ergodic distribution of a single firm's states, so the macro response is the average of micro marginal responses.

Note that Lemma B.3 requires no continuity of f_h in ϵ and no density for S, $\mathcal{D}$ or ϵ_{it}.[24]

[24]This is because it differentiates nothing inside an expectation: it identifies R_h with the cross-sectional average of the shifted price path and then differentiates that average. Whether R_h is differentiable at 0 is a property of the model's aggregate response, satisfied under the assumptions of Section 2.1 on the idiosyncratic shocks. Take the case (MC): conditional on $(S, \epsilon_{it}, \mathcal{D})$, a firm's price at $t+h$ is a step function of the

Step (a) is also where the limits of the identification appear: the regression generates the displacement whose profile is the impulse response of p^* to the shifter innovation, a permanent step for a random-walk shifter. An mc path whose profile differs from the shifter's own impulse response—a news shock, or a transitory shift when the shifter is a random walk—is not that displacement, the policies respond to the anticipated path, and no regression on innovations to that shifter can reproduce it. This is why the bridge delivers exactly $J^m_{P,mc}\mathbf{1}$ and no other feature of the matrix.

Persistent shifters. If $\vartheta_{it} = \rho\vartheta_{i,t-1} + \epsilon_{it}$, a shift δ in ϵ_{it} raises $p^*_{i,t+s}$ by $\omega\rho^s\delta$, other innovations held fixed, exactly as an unanticipated cost path $(\delta, \rho\delta, \rho^2\delta, \dots)$ does with the price level fixed. The firm's state now includes ϑ_{it}; along the shifted path the pair (gap, ϑ) evolves exactly as it would under an unanticipated cost path $(\delta, \rho\delta, \rho^2\delta, \dots)$ starting from the same state, with identical policy functions (the policy under the anticipated cost path is formally time-varying but coincides at $t + j$ with the stationary policy evaluated at $\vartheta_{i,t+j} + \rho^j\delta$). Lemma B.3, combined with Lemma B.4, then gives $b^m_h = [J^m_{P,mc}\tilde{\boldsymbol{\vartheta}}]_h$ with $\tilde{\boldsymbol{\vartheta}} = (1, \rho, \rho^2, \dots)'$, provided ϵ_{it} is independent of $(x_{i,t-1}, \vartheta_{i,t-1})$, the domination (B.46) still holds (it does, since $\rho \leq 1$), and the same ρ is used for the shifter and for the cost path.

B.5.2. Step 2: The Average Derivative Is the Regression Coefficient

Since ϵ_{it} is independent of $(S, \mathcal{D})$, by definition of the OLS coefficient,

$$b^m_h = \frac{\mathbb{E}_{S,\epsilon,\mathcal{D}}[\, f_h(\epsilon_{it}; S, \mathcal{D})\, \epsilon_{it}]}{\sigma^2_\vartheta} = \mathbb{E}_{S,\mathcal{D}}\left[\frac{\mathbb{E}_\epsilon[\, f_h(\epsilon; S, \mathcal{D})\, \epsilon]}{\sigma^2_\vartheta}\right]. \tag{B.39}$$

The lemma gives two sufficient conditions under which this regression coefficient equals A_h.

---

common shift: it equals the old price until δ tips the firm across the band and the reset price afterwards. Its pointwise derivative exists for almost every realization and takes values in $\{0, \omega\}$ (ω if the firm has adjusted since t, zero otherwise), but its expectation is only the intensive margin: the integrand jumps at the values of δ at which the firm's adjustment decision flips, and the mass swept across those jumps, the extensive margin, is missed by any argument that differentiates realization by realization. A derivative of the *aggregate* exists because, across firms, the pre-adjustment gap $x_{i,t-1} - \omega(\epsilon^z_{it} + \epsilon_{it})$ has a continuous density, so that the mass tipped over by a small δ is proportional to δ.

LEMMA B.4 (Regression coefficient equals average derivative). *Let $(S, \epsilon_{it}, \mathcal{D})$ be mutually independent, with ϵ_{it} of mean zero and variance σ^2_ϑ, and let $|f_h(\epsilon; S, \mathcal{D})| \leq M_h(S, \mathcal{D}) + |\epsilon|$ for all ϵ, with $\mathbb{E}_{S,\mathcal{D}}[M_h] < \infty$, so that b^m_h is well defined. Suppose that either*

(G) *$\epsilon_{it} \sim N(0, \sigma^2_\vartheta)$; or*

(L) *$F_h(\epsilon; S) \equiv \mathbb{E}_{\mathcal{D}}[f_h(\epsilon; S, \mathcal{D})]$ is affine in ϵ on an open set containing the support of ϵ_{it}, for μ_S-almost every S: $F_h(\epsilon; S) = a_h(S) + s_h(S)\,\epsilon$ there, with $\mathbb{E}_S|a_h|, \mathbb{E}_S|s_h| < \infty$.*

Then A_h exists and $A_h = b^m_h$. Under (G), moreover, $A_h = \mathbb{E}_{S,\mathcal{D}}\big[\frac{d}{d\delta}\,\mathbb{E}_\epsilon[f_h(\epsilon + \delta; S, \mathcal{D})]\big|_{\delta=0}\big]$.

PROOF. *Case* (L). Because $(S, \epsilon_{it}, \mathcal{D})$ are mutually independent and f_h is dominated by $M_h + |\epsilon|$, Fubini's theorem allows integrating $\mathcal{D}$ out first in both objects, for $|\delta|$ small enough that $\epsilon + \delta$ stays in the set where F_h is affine. For the regression coefficient, $\mathbb{E}_\epsilon[F_h(\epsilon; S)\,\epsilon] = \mathbb{E}_\epsilon[(a_h + s_h\epsilon)\epsilon] = s_h(S)\,\sigma^2_\vartheta$, so $b^m_h = \mathbb{E}_S[s_h]$. For the average derivative, $\mathbb{E}_{S,\epsilon,\mathcal{D}}[f_h(\epsilon + \delta)] = \mathbb{E}_S\mathbb{E}_\epsilon[a_h + s_h(\epsilon + \delta)] = \mathbb{E}_S[a_h] + \delta\,\mathbb{E}_S[s_h]$, which is affine in δ, so $A_h = \mathbb{E}_S[s_h]$. No exchange of limit and expectation is needed, and no property of the law of ϵ_{it} beyond a finite variance and the domination above.

Case (G). The argument has three parts: Stein's identity for a single Gaussian variable, its application conditionally on $(S, \mathcal{D})$, and the exchange of $d/d\delta$ with $\mathbb{E}_{S,\mathcal{D}}$.

Stein's identity. Let $\epsilon \sim N(0, \sigma^2)$ with density φ_σ, let f be measurable with polynomial growth, $|f(u)| \leq C_f(1 + |u|^q)$ for some constants C_f, q, and fix $c > 0$ (constants local to this appendix, unrelated to the bounds m, c of (A2)). Then

$$\frac{d}{d\delta}\,\mathbb{E}_\epsilon\big[f(\epsilon + \delta)\big]\Big|_{\delta=0} = \frac{\mathbb{E}_\epsilon[f(\epsilon)\,\epsilon]}{\sigma^2}. \tag{B.40}$$

This is Stein's lemma (Stein 1981, Lemma 1) in location-derivative form. The classical statement, $\mathbb{E}_\epsilon[f(\epsilon)\epsilon] = \sigma^2\,\mathbb{E}_\epsilon[f'(\epsilon)]$, assumes f absolutely continuous, which the menu cost response is not; the form (B.40) needs no regularity of f because the derivative falls on the density. Writing $\mathbb{E}_\epsilon f(\epsilon + \delta) = \int f(u)\varphi_\sigma(u - \delta)\,du$, the integrand is C^∞ in δ with derivative $-f(u)\varphi'_\sigma(u - \delta) = f(u)\,\frac{u-\delta}{\sigma^2}\,\varphi_\sigma(u - \delta)$, which for $|\delta| \leq c$ is dominated, uniformly in δ, by

$$G(u) \equiv \frac{C_f\,(1 + |u|^q)(|u| + c)}{\sigma^3\sqrt{2\pi}}\,e^{-[(|u|-c)_+]^2/2\sigma^2}, \qquad \int G(u)\,du < \infty, \tag{B.41}$$

so differentiation under the integral sign is licensed (mean value theorem on the difference quotients, then dominated convergence), and the change of variable $u = \epsilon + \delta$

gives $\frac{d}{d\delta}\mathbb{E}_\epsilon f(\epsilon+\delta) = \mathbb{E}_\epsilon[f(\epsilon+\delta)\,\epsilon]/\sigma^2$ for all $|\delta|< c$. When f is piecewise C^1 with finitely many jumps $\Delta f_j \equiv f(u_j^+) - f(u_j^-)$ at points u_j and $|f'|\leq C_f(1+|u|^q)$ on the smooth stretches, integrating by parts stretch by stretch gives the "intensive plus extensive" form $\mathbb{E}_\epsilon[f'(\epsilon)] + \sum_j \Delta f_j\,\varphi_\sigma(u_j)$, which is the classical statement when there are no jumps. Conversely, (B.40) characterizes the Gaussian (Stein 1972): it holds for every f only if $\varphi' = -(u/\sigma^2)\varphi$; for linear f it holds for any distribution.

Conditional application. The domination $|f_h(u;S,\mathcal{D})|\leq M_h(S,\mathcal{D})+|u|$ gives $f_h(\cdot\,;S,\mathcal{D})$ polynomial growth of degree one for every $(S,\mathcal{D})$, so under (G) (B.40) applies conditionally on $(S,\mathcal{D})$:

$$b_h^m = \mathbb{E}_{S,\mathcal{D}}\Big[\frac{d}{d\delta}\,\mathbb{E}_\epsilon\big[f_h(\epsilon+\delta;S,\mathcal{D})\big]\Big|_{\delta=0}\Big]. \tag{B.42}$$

Exchanging the derivative and the expectation over $(S,\mathcal{D})$**.** Write $\gamma_h(\delta;S,\mathcal{D}) \equiv \mathbb{E}_\epsilon[f_h(\epsilon+\delta;S,\mathcal{D})]$, so that (B.42) reads $b_h^m = \mathbb{E}_{S,\mathcal{D}}[\gamma_h'(0)]$, while the claim is $b_h^m = \frac{d}{d\delta}\mathbb{E}_{S,\mathcal{D}}[\gamma_h(\delta)]\big|_{\delta=0}$. Three facts let the derivative move outside the expectation.

(1) For every $(S,\mathcal{D})$, γ_h is a smooth function of δ, because δ enters $\gamma_h(\delta) = \int f_h(u;S,\mathcal{D})\,\varphi_\sigma(u-\delta)\,du$ only through the Gaussian density (by independence of ϵ_{it} from $(S,\mathcal{D})$); by the computation behind (B.40), $\gamma_h'(\delta) = \mathbb{E}_\epsilon[f_h(\epsilon+\delta;S,\mathcal{D})\,\epsilon]/\sigma_\vartheta^2$.

(2) Its derivative is bounded, uniformly over $|\delta|\leq c$, by an integrable function G_h of $(S,\mathcal{D})$ alone.[25]

(3) Dominated convergence then applies to the difference quotients: by the mean value theorem, $\{\gamma_h(\delta+r)-\gamma_h(\delta)\}/r = \gamma_h'(\tilde\delta)$ for some $\tilde\delta$ between δ and $\delta+r$, so the quotients are bounded by G_h whenever $|\delta|+|r|\leq c$, and letting $r\to 0$ inside the expectation gives $\frac{d}{d\delta}\mathbb{E}_{S,\mathcal{D}}[\gamma_h(\delta)] = \mathbb{E}_{S,\mathcal{D}}[\gamma_h'(\delta)]$ on $(-c,c)$.

[25] By the domination hypothesis, $|f_h(\epsilon+\delta;S,\mathcal{D})|\leq M_h+|\epsilon|+c$ for $|\delta|\leq c$, so

$$|\gamma_h'(\delta;S,\mathcal{D})| \leq \frac{\mathbb{E}_\epsilon[(M_h+|\epsilon|+c)\,|\epsilon|]}{\sigma_\vartheta^2} = \frac{(M_h+c)\sqrt{2/\pi}}{\sigma_\vartheta} + 1 \equiv G_h(S,\mathcal{D}), \tag{B.43}$$

and $\mathbb{E}_{S,\mathcal{D}}[G_h]<\infty$ because $\mathbb{E}_{S,\mathcal{D}}[M_h]<\infty$. The same domination gives $\mathbb{E}_{S,\mathcal{D}}|\gamma_h(\delta)|\leq \mathbb{E}_{S,\mathcal{D}}[M_h]+\mathbb{E}_\epsilon|\epsilon|+c<\infty$, so $\mathbb{E}_{S,\mathcal{D}}[\gamma_h(\delta)]$ is well defined for $|\delta|\leq c$. Note that $\mathbb{E}_{S,\mathcal{D}}[G_h]<\infty$ requires finite first moments under $\mu_S\otimes\mu_\mathcal{D}$: $\mathbb{E}_S|x_{i,t-1}|<\infty$ and finite $\mathbb{E}_\mathcal{D}|\epsilon_{it}^z|$, $\mathbb{E}_\mathcal{D}|\epsilon_{i,t+j}^z|$, $\mathbb{E}_\mathcal{D}|\epsilon_{i,t+j}|$ for $j\leq h$ but the menu-cost and lottery draws do not enter M_h.

At $\delta = 0$, (B.42) therefore becomes

$$b_h^m = \frac{d}{d\delta}\,\mathbb{E}_{S,\epsilon,\mathcal{D}}[f_h(\epsilon_{it}+\delta; S, \mathcal{D})]\Big|_{\delta=0}. \tag{B.44}$$

The right-hand side is A_h, so $b_h^m = A_h$; and (B.42) together with (B.44) is the second claim of the lemma.

□

Note that no density, continuity or stationarity of μ_S or $\mu_{\mathcal{D}}$ is used here. In particular, μ_S has an atom at the reset gap $x^* = 0$, where every adjuster sits, and this is harmless because the smoothing in δ is supplied by ϵ_{it} alone.

The three models. *Model* (C). Conditional on the lottery draws, which belong to $\mathcal{D}$, f_h is affine in ϵ with slope $\omega\,\mathbf{1}\{\text{a reset occurs in } [t, t+h]\}$: under the random walk every reset price (22) moves by ω per unit of ϵ_{it}, and a price that has not been reset does not move. Averaging over the draws, $F_h(\epsilon; S) = \mathbb{E}_{\mathcal{D}}[f_h]$ is affine in ϵ with slope $\omega \Pr(\text{reset in } [t, t+h]) = \omega(1-\alpha^{h+1})$ (equation (23)), and the domination $|f_h| \leq M_h + |\epsilon|$ holds by the representation (B.45) below with the reset price in place of p^*_{i,τ^*} (plus the forecast terms of z, of finite mean, when z is not a random walk): condition (L) holds for any distribution of ϵ, and Lemma B.4 gives $A_h = b_h^{\mathrm{C}} = \omega(1-\alpha^{h+1})$.

Model (GL). Conditional on hits and turbulence draws, f_h is not linear in ϵ (whether the firm adjusts depends on ϵ through its gap), but after integrating over the turbulence draws it is: Lemma B.2 holds for every shifter path under (A2)–(A3), so $F_h(\epsilon; S) = \mathbb{E}_{\mathcal{D}}[f_h]$ is affine in ϵ with slope $\omega(1-\eta_\lambda^{h+1})$ (Appendix B.3.2): condition (L) holds to the order of the GL approximation, for any distribution of ϵ with support compatible with (A2). Affinity fails only on the set of histories whose last hit is older than the horizon over which (A2) bounds the cumulated shifter, a set of μ_S-mass $O(m)$ on which the conditional mean of $x_{i,t+h}$ is bounded by $\bar{x}$, so the error in b_h is $O(m\bar{x})$, the order of the approximation in the reset rule (A.6); the domination $|f_h| \leq M_h + |\epsilon|$ holds as in (C). Lemma B.4 then gives $A_h = b_h^{\mathrm{GL}} = \omega(1-\eta_\lambda^{h+1}) = \omega[J^{\mathrm{C}}_{P,p^*}(\eta_\lambda)\mathbf{1}]_h$ to that order. Note that (A2) is a bounded-support condition on ϵ, incompatible with the Gaussian shifter of (MC): the two models place different conditions on the identifying shock, and thin-tailed shifters make the GL error exponentially small (see Section B.5.4, Scope).

Model (*MC*), *Gaussian* ϵ. Here f_h is not differentiable in ϵ: it jumps whenever ϵ pushes the pre-adjustment gap $x_{i,t-1} - \omega(\epsilon^z_{it} + \epsilon)$ across the band at t or, for a firm that has not yet adjusted, at a later date. Condition (G) is verified from the policy. If the firm does not adjust between t and $t+h$ then $f_h = 0$. Otherwise let $\tau^* \in \{t, \dots, t+h\}$ be its last adjustment date; since the reset gap is $x^* = 0$, $p_{i,t+h} = p^*_{i,\tau^*}$ and, with the price level fixed,

$$f_h(\epsilon; S, \mathcal{D}) = p^*_{i,\tau^*} - p_{i,t-1} = -x_{i,t-1} + \omega(\epsilon^z_{it} + \epsilon) + \sum_{j=1}^{\tau^*-t} v_{i,t+j}, \qquad v_{i\tau} \equiv \omega(\epsilon^z_{i\tau} + \epsilon_{i\tau}). \tag{B.45}$$

Hence, for every ϵ and every $|\delta| \leq c$, using $\omega \leq 1$,

$$\begin{aligned} \left| f_h(\epsilon + \delta; S, \mathcal{D}) \right| &\leq M_h(S, \mathcal{D}) + \omega(|\epsilon| + c) \leq M_h(S, \mathcal{D}) + |\epsilon| + c, \\ M_h &\equiv |x_{i,t-1}| + \omega|\epsilon^z_{it}| + \sum_{j=1}^{h} |v_{i,t+j}|, \end{aligned} \tag{B.46}$$

and $\mathbb{E}_{S,\mathcal{D}}[M_h] < \infty$: post-adjustment gaps lie in the band, so $|x_{i,t-1}| \leq \bar{x}$ (with a menu-cost distribution H of unbounded support, $\mathbb{E}_S|x_{i,t-1}| < \infty$ is the assumption of Proposition 1; it is automatic when H has bounded support), and the innovations are Gaussian, so that the domination hypothesis holds and (G) holds.

Proof of Proposition 1. Under the conditions of the proposition the hypotheses of Lemma B.4 hold: the mutual independence of $(S, \epsilon_{it}, \mathcal{D})$ and the domination of f_h in every model, and (L) in (C), (L) to the order of the approximation in (GL) under (A2)–(A3), (G) in (MC) with Gaussian ϵ_{it}. Lemma B.4 therefore shows that A_h exists and equals b^m_h. By (B.36), the existence of A_h is the differentiability of R_h at 0 required in Lemma B.3, which then gives $[J^m_{P,mc}\mathbf{1}]_h = A_h = b^m_h$. □

B.5.3. Intuition for Step 2, Case (G): The Regression as a Weighted Average Derivative

This section rereads the proof of case (G) of Lemma B.4. The regression coefficient and the average derivative are both averages of the same *local responses* of the price change to the shock, under two different weightings; they agree for every response function exactly when the two weightings coincide, and that happens only for the Gaussian. The statements below are written for a discontinuous response, since that is the menu cost case.

Throughout, $(S, \mathcal{D})$ is fixed, $f \equiv f_h(\cdot\,; S, \mathcal{D})$, and ϵ is a single random variable with law μ, mean zero and variance $\sigma^2 \in (0, \infty)$; every expectation is $\mathbb{E}_\epsilon$. The shock is measured in desired-price units, $\omega\epsilon$, so that $\omega = 1$ throughout this subsection without loss of generality.

Local responses. In model (MC), f is affine on each of finitely many intervals, with slopes in $\{0, 1\}$ and at most $2(h+1)$ jump points, and $|f(u)| \leq M_h + |u|$ by (B.46). The reason is the Ss structure: given the draws in $\mathcal{D}$, the set of shocks u under which the firm has not adjusted by a given date is an intersection of intervals, hence an interval, and these intervals shrink with the date; so the first-adjustment date is piecewise constant in u, with at most two boundary points per date, and once the firm has adjusted its gap is reset to $x^* = 0$ and nothing later depends on u, so that f is affine with slope 1 on each piece where the firm adjusts by $t + h$ and equal to 0 where it does not. We take f right-continuous at its jump points $u_1 < \cdots < u_J$, a normalization that is immaterial when μ has no atom there, and write $\Delta f_j \equiv f(u_j^+) - f(u_j^-)$. The *local response* at u is the slope $f'(u)$ off the jump points and the jump Δf_j at u_j; to average local responses against a weight k we use the generalized derivative $Df \equiv f'(u)\,du + \sum_j \Delta f_j\, \delta_{u_j}$, i.e.

$$\int k\, dDf \equiv \int f'(u)\, k(u)\, du + \sum_{j=1}^{J} \Delta f_j\, k(u_j), \tag{B.47}$$

so that jumps count as point masses, and the fundamental theorem of calculus reads

$$f(b) - f(a) = Df((a, b]) = \int_a^b f'(u)\, du + \sum_{j:\, a < u_j \leq b} \Delta f_j, \qquad a < b. \tag{B.48}$$

For the impact response of a firm that starts at a zero gap, $f(u) = u\,\mathbf{1}\{|u| > \bar{x}\}$, one has $f' = \mathbf{1}\{|u| > \bar{x}\}$ and $\Delta f = \bar{x}$ at both $u = \pm\bar{x}$, so

$$\int k\, dDf = \int_{|u| > \bar{x}} k(u)\, du + \bar{x}\,[k(\bar{x}) + k(-\bar{x})]\,: \tag{B.49}$$

with k the density of the shock, the intensive margin (the adjusters move one for one with the desired price) plus the extensive margin (the firms swept across the thresholds, each moving by $\bar{x}$), the structure of the impact entry (B.17) and of b_0^{MC}.

The two weightings. The following proposition shows that the regression coefficient and the average derivative are both averages of the same local responses $f'(u)$, under two different weightings. They agree for every response function exactly when the two weightings coincide, and that happens only for the Gaussian.

PROPOSITION B.2 (Two weightings of the local responses). *Let f be piecewise C^1 with finitely many jumps and $|f(u)|+|f'(u)|\le C_f(1+|u|)$ off the jump points, and let μ have mean zero, variance σ^2, $\mathbb{E}_\epsilon|\epsilon|^3<\infty$, and no atom at the jump points of f.*

(*i*) Regression weighting. *Define*

$$w(u) \equiv \frac{1}{\sigma^2}\,\mathbb{E}_\epsilon[\epsilon\,\mathbf{1}\{\epsilon > u\}] = \frac{1}{\sigma^2}\,\mathbb{E}_\epsilon[(-\epsilon)\,\mathbf{1}\{\epsilon \le u\}], \qquad u \in \mathbb{R}, \tag{B.50}$$

the two forms being equal because $\mathbb{E}_\epsilon\epsilon = 0$. Then w is a probability density, without atoms, continuous except at an atom of μ located at $u \neq 0$, where it jumps by $u\,\mu(\{u\})/\sigma^2$, and

$$\frac{\mathbb{E}_\epsilon[f(\epsilon)\,\epsilon]}{\sigma^2} = \int w\,dDf = \int f'(u)\,w(u)\,du + \sum_{j=1}^{J} \Delta f_j\,w(u_j). \tag{B.51}$$

(*ii*) Shift weighting. *If μ has a continuously differentiable density φ with $|u|\varphi(u)\to 0$ as $|u|\to\infty$ and $\int \sup_{|\delta|\le c}(1+|u|)\,|\varphi'(u-\delta)|\,du<\infty$ for some $c>0$, then*

$$\frac{d}{d\delta}\,\mathbb{E}_\epsilon[f(\epsilon+\delta)]\Big|_{\delta=0} = -\int f(u)\,\varphi'(u)\,du = \int \varphi\,dDf. \tag{B.52}$$

(*iii*) When they coincide. *If μ has a continuous density φ, then w is continuously differentiable with*

$$w'(u) = -\frac{u\,\varphi(u)}{\sigma^2}, \tag{B.53}$$

and $w=\varphi$ on $\mathbb{R}$ if and only if $\varphi=\varphi_\sigma$, the $N(0,\sigma^2)$ density.

PROOF. (i) Fix a continuity point c of f. Since $\mathbb{E}_\epsilon\epsilon=0$, $\mathbb{E}_\epsilon[f(\epsilon)\epsilon]=\mathbb{E}_\epsilon[(f(\epsilon)-f(c))\epsilon]$, and

by (B.48)

$$f(\epsilon)-f(c)=\begin{cases}\displaystyle\int_c^{\infty} f'(u)\,\mathbf{1}\{u<\epsilon\}\,du+\sum_j \Delta f_j\,\mathbf{1}\{c<u_j\le\epsilon\}, & \epsilon>c,\\ \displaystyle-\int_{-\infty}^{c} f'(u)\,\mathbf{1}\{\epsilon<u\}\,du-\sum_j \Delta f_j\,\mathbf{1}\{\epsilon<u_j\le c\}, & \epsilon\le c.\end{cases}\tag{B.54}$$

Multiply by ϵ and integrate against μ. Fubini's theorem allows the order of integration in the du terms to be exchanged, the integrand being bounded by $C_f(1+|u|)|\epsilon|$ on a u-range of length at most $|\epsilon|+|c|$ on which $|u|\le|\epsilon|+|c|$, which is integrable by the third moment; the jump terms are finite sums. Reading off $\mathbb{E}_\epsilon[\epsilon\mathbf{1}\{\epsilon>u\}]=\sigma^2 w(u)$ and $\mathbb{E}_\epsilon[(-\epsilon)\mathbf{1}\{\epsilon\le u\}]=\sigma^2 w(u)$ from (B.50), and using $\mu(\{u_j\})=0$ so that $\mathbf{1}\{u_j\le\epsilon\}=\mathbf{1}\{u_j<\epsilon\}$ almost surely,

$$\mathbb{E}_\epsilon[f(\epsilon)\,\epsilon]=\sigma^2\Big[\int_{-\infty}^{\infty} f'(u)\,w(u)\,du+\sum_{j=1}^{J}\Delta f_j\,w(u_j)\Big],\tag{B.55}$$

which is (B.51); the base point c drops out, as it must. As for w: the first form in (B.50) shows $w\ge 0$ for $u\ge 0$ and the second for $u<0$; Tonelli's theorem gives $\int w=\sigma^{-2}\mathbb{E}_\epsilon[\epsilon^2]=1$; and dominated convergence on the indicators gives right-continuity and $w(u^-)-w(u)=u\,\mu(\{u\})/\sigma^2$.

(ii) The first equality is the argument behind (B.40)–(B.41) with φ in place of φ_σ: the integrand $f(u)\varphi(u-\delta)$ of $\mathbb{E}_\epsilon f(\epsilon+\delta)$ is C^1 in δ, its derivative is dominated for $|\delta|\le c$ by $C_f(1+|u|)\sup_{|\delta'|\le c}|\varphi'(u-\delta')|$, which the stated condition makes integrable (it is what (B.41) delivers in the Gaussian case), and the mean value theorem with dominated convergence lets the derivative pass under the integral. For the second, integrate $-\int f\varphi'$ by parts on each stretch between consecutive jump points: the two boundary terms at a jump point combine, by continuity of φ, into $f(u_j^+)\varphi(u_j)-f(u_j^-)\varphi(u_j)=\Delta f_j\varphi(u_j)$, the terms at $\pm\infty$ vanish because $|f\varphi|\le C_f(1+|u|)\varphi\to 0$, and $\int|f'|\varphi<\infty$ by the first moment.

(iii) (B.53) is the fundamental theorem of calculus applied to $w(u)=\sigma^{-2}\int_u^\infty v\varphi(v)\,dv$. If $w=\varphi$ then $\varphi'=-u\varphi/\sigma^2$, a linear first-order equation whose solutions are $\varphi(0)e^{-u^2/2\sigma^2}$, and $\int\varphi=1$ fixes $\varphi(0)=(\sigma\sqrt{2\pi})^{-1}$. Conversely $v\varphi_\sigma(v)=-\sigma^2\varphi_\sigma'(v)$ gives $\int_u^\infty v\varphi_\sigma=\sigma^2\varphi_\sigma(u)$. □

Interpretation of the two weightings: Chords and tangents. The proof of Proposition B.2(i) reads the regression through the total displacement $f(\epsilon) - f(c)$, the sum of the local responses along the path from c to ϵ; the local response at u is credited to every firm whose path crosses u, each carrying its regression weight ϵ, which is why the weight at u is the shock-weighted mass of firms beyond u, (B.50). Part (ii) of the proposition instead credits the local response at u to the firms located at u, with weight $\varphi(u)$. The regression measures *chord* slopes between c and ϵ, the economy responds with *tangent* slopes at ϵ, and w records where along the chords the regression samples the tangents. Equivalently, by the Tonelli computation in that proof with $c = 0$, w is the density of a point drawn uniformly on the segment between 0 and ϵ in a sample in which firms are drawn with probability proportional to ϵ^2: once because longer paths contain more points, once because the regression weights each firm by ϵ. Part (iii) says that for the bell curve, and only for it, the shock-weighted mass beyond u equals σ^2 times the density at u, so chord sampling and endpoint sampling agree at every u; this is the same property that makes the score linear, $\varphi'_\sigma = -(u/\sigma^2)\varphi_\sigma$.

If f' is the same number everywhere and f has no jumps, $\int w\, dDf = \int \varphi\, dDf$ whatever w and φ are: the weighting is irrelevant and any law works. Applied to $F_h = \mathbb{E}_{\mathcal{D}} f_h$, this is the content of case (L) of Lemma B.4, which covers the Calvo model and, to the order of its approximation, the GL model.

Note: A general fact about regressions on a randomized continuous treatment. Nothing in Proposition B.2 is specific to price setting. Consider any outcome $y = g(\epsilon; \zeta)$ of a continuous treatment ϵ that is randomly assigned, i.e. independent of the unit's characteristics ζ, with response functions $g(\cdot\,; \zeta)$ that may be heterogeneous across units and nonlinear in the treatment, and let $m(u) \equiv \mathbb{E}_\zeta[g(u; \zeta)]$ be the average response, assumed piecewise C^1 with finitely many jumps and linear growth. The average marginal treatment effect is $\int \varphi\, dDm$, which equals $\mathbb{E}_\epsilon[m'(\epsilon)]$ when m is smooth and includes the mass swept across the jumps when individual responses are discontinuous, as in part (ii). The OLS coefficient of y on ϵ is $\mathbb{E}[y\epsilon]/\sigma^2 = \mathbb{E}_\epsilon[m(\epsilon)\epsilon]/\sigma^2 = \int w\, dDm$ by independence and part (i): heterogeneity enters only through m. The two coincide for a given m when m is affine, which covers heterogeneous but linear responses, and for every m if and only if the treatment is Gaussian (part (iii)). Otherwise OLS is a weighted average of marginal effects

with the weights w, a result due to Yitzhaki (1996) and discussed by Angrist and Krueger (1999).

B.5.4. Non-Gaussian Identifying Shocks

B.5.5. General Case

Let $\bar{F}_h(\epsilon) \equiv \mathbb{E}_{S,\mathcal{D}}[f_h(\epsilon; S, \mathcal{D})]$ be the response averaged over the state and the other draws. It is smooth even though each realization f_h jumps, because the driving innovations ϵ^z have smooth densities and spread the jump points out; this is the only place where a property of ϵ^z beyond finite first moments is used. By Proposition B.2, the regression averages the local responses of $\bar{F}_h$ under the weight w and the shift experiment averages them under the density of the shock, and the two agree for every response only when the shock is Gaussian. Their moments differ otherwise: $\int u\, w(u)\, du = \mu_3/(2\sigma_\vartheta^2)$ and $\int u^2 w(u)\, du = \sigma_\vartheta^2\, \text{kurt}/3$, so a skewed shock tilts the regression weight and a fat-tailed one spreads it beyond the population density.

The bias. Suppose ϵ has mean zero, variance σ_ϑ^2 and bounded absolute standardized moments up to order five, and that $\bar{F}_h$ is C^4 near zero with linear growth. Expanding $\bar{F}_h$ inside the two averages,

$$\frac{\mathbb{E}_\epsilon[\bar{F}_h(\epsilon)\,\epsilon]}{\sigma_\vartheta^2} - \frac{d}{d\delta}\,\mathbb{E}_\epsilon[\bar{F}_h(\epsilon+\delta)]\Big|_{\delta=0} = \frac{\bar{F}_h''(0)}{2}\,\text{skew}\,\sigma_\vartheta + \frac{\bar{F}_h'''(0)}{6}\,(\text{kurt}-3)\,\sigma_\vartheta^2 + O(\sigma_\vartheta^3), \tag{B.56}$$

which vanishes at all orders in the Gaussian case by (B.40). At the zero-inflation steady state the model is invariant to the sign flip of all gaps and innovations, so $\bar{F}_h$ is odd up to $O(\sigma_\vartheta^3)$ so the leading bias is $\frac{1}{6}\bar{F}_h'''(0)(\text{kurt}-3)\sigma_\vartheta^2$, proportional to excess kurtosis.

Solution: regress on the score. Let p be the density of ϵ, absolutely continuous with $\int \sup_{|\delta|\le c}(1+|u|)\,|p'(u-\delta)|\,du < \infty$ for some $c > 0$ (Gaussian, Student-t and Laplace densities qualify), and let $s(\epsilon) \equiv -p'(\epsilon)/p(\epsilon)$ be its score. Integration by parts, with jumps counted as in (B.47), gives $\mathbb{E}_\epsilon[f(\epsilon)s(\epsilon)] = \frac{d}{d\delta}\mathbb{E}_\epsilon[f(\epsilon+\delta)]|_{\delta=0}$ for every piecewise-C^1 f of linear growth; the Gaussian score is $\epsilon/\sigma_\vartheta^2$, which recovers (B.40). Repeating the two steps

of the proof with $s(\epsilon_{it})$ in place of $\epsilon_{it}/\sigma_\vartheta^2$,

$$\mathbb{E}_{S,\epsilon,\mathcal{D}}[(p_{i,t+h} - p_{i,t-1})\, s(\epsilon_{it})] = [J_{P,mc}^m \mathbf{1}]_h \tag{B.57}$$

exactly, for any such density. Since $\mathbb{E}_\epsilon[\epsilon\, s(\epsilon)] = 1$ and $\mathbb{E}_\epsilon[s(\epsilon)] = 0$, this is the coefficient of an IV regression of price changes on ϵ_{it} with instrument $s(\epsilon_{it})$: the average-derivative estimator of Stoker (1986) and Härdle and Stoker (1989). p can be estimated from the observed panel of ϵ_{it}.

B.5.6. The Turbulence Regression in the Gertler–Leahy Model

Suppose the econometrician regresses price changes on the turbulence innovation ϵ_{it}^z itself (equal to zero for firms not hit), a violation of the conditions of Proposition 1 in (GL). This section shows that the resulting coefficients are $b_h^z = \omega(1 - \iota^{\mathrm{GL}}\, \eta_\lambda^h)$ with $\iota^{\mathrm{GL}} \equiv (1 - \lambda)(2\bar{x}/\varsigma)^3 < 1/8$: the impact pass-through is close to its long-run value ω and $b_1^z/b_0^z - 1$ does not recover η_λ, because conditional on a nonzero innovation the firm has been hit, and hit firms pass the shock through fully on average. Set mc and P constant; the shifter innovations are independent of ϵ_{it}^z and drop out of every covariance below, so we also set ϑ constant, and $p_{it}^* = \omega z_{it}$ up to a constant and $p_{i,t+h}^* - p_{i,t-1}^* = \omega(z_{i,t+h} - z_{i,t-1})$ is a random walk in the turbulence innovations. Define

$$b_h^z \equiv \frac{\mathrm{Cov}(p_{i,t+h} - p_{i,t-1}, \epsilon_{it}^z)}{\mathrm{Var}(\epsilon_{it}^z)}$$

over all firms, with $\epsilon_{it}^z = 0$ for firms not hit at t; since those firms contribute zero to both the covariance and the variance, b_h^z equals the same ratio computed among hit firms. Corollary 1 applies with ϵ^z in place of ϵ: $b_h^z = \omega + \mathrm{Cov}(x_{i,t+h}, \epsilon_{it}^z)/\mathrm{Var}(\epsilon_{it}^z)$. It is convenient to work with the turbulence innovation in desired-price units, $u_{it}^z \equiv \omega\epsilon_{it}^z$, uniform on $[-\varsigma/2, \varsigma/2]$ at a hit; since $\epsilon^z = u^z/\omega$,

$$\frac{\mathrm{Cov}(x_{i,t+h}, \epsilon_{it}^z)}{\mathrm{Var}(\epsilon_{it}^z)} = \omega\, \frac{\mathrm{Cov}(x_{i,t+h}, u_{it}^z)}{\mathrm{Var}(u_{it}^z)}.$$

With mc, P and ϑ constant the reset price net of the turbulence component, $p_{it}^{\mathrm{new}} - \omega z_{it}$, is a constant, normalized to zero. Consider a firm hit at t with pre-hit gap $x_- \equiv p_{i,t-1} - \omega z_{i,t-1} \in [-\bar{x}, \bar{x}]$, independent of u_{it}^z. Its would-be gap is $y = x_- - u_{it}^z$, and its gap after the hit is $x_{it} = y\,\mathbf{1}\{|y| \leq \bar{x}\}$ if it has no free draw (probability $1 - \lambda$) and 0 if it has one.

Substituting $u = -y$, which ranges over an interval of width ς containing $[-\bar{x}, \bar{x}]$ by (A3),

$$\text{(B.58)} \quad \mathbb{E}[x_{it}\, u^z_{it} \mid x_-] = \mathbb{E}[y\,(x_- - y)\,\mathbf{1}\{|y| \leq \bar{x}\} \mid x_-] = -\frac{1}{\varsigma}\int_{-\bar{x}}^{\bar{x}} u\,(u + x_-)\,du = -\frac{2\bar{x}^3}{3\varsigma},$$

independent of x_-, while $\text{Var}(u^z_{it} \mid \text{hit}) = \varsigma^2/12$. Multiplying by the probability $1 - \lambda$ of no free draw, $\text{Cov}(x_{it}, u^z_{it})/\text{Var}(u^z_{it}) = -(1-\lambda)\,8\bar{x}^3/\varsigma^3 = -\iota^{\text{GL}}$ and $b^z_0 = \omega(1 - \iota^{\text{GL}})$. For later horizons, $x_{i,t+1} = x_{it}$ with probability η_λ; with probability λ the firm resets freely, so $x_{i,t+1} = 0$; and with the remaining probability it is hit again with would-be gap $x_{it} - u^z_{i,t+1}$, uniform on an interval containing the band, so that $\mathbb{E}[x_{i,t+1} \mid x_{it}, \text{hit}] = \varsigma^{-1}\int_{-\bar{x}}^{\bar{x}} u\,du = 0$. Therefore $\mathbb{E}[x_{i,t+h} \mid x_{it}] = \eta^h_\lambda x_{it}$, $\text{Cov}(x_{i,t+h}, u^z_{it}) = \eta^h_\lambda\,\text{Cov}(x_{it}, u^z_{it})$, and

$$\text{(B.59)} \qquad b^z_h = \omega(1 - \iota^{\text{GL}}\,\eta^h_\lambda),$$

the result stated above. In particular $b^z_h \to \omega$ (full long-run pass-through of the desired price) and $b^z_1/b^z_0 - 1 = \iota^{\text{GL}}(1-\eta_\lambda)/(1-\iota^{\text{GL}})$. Contrast with the shifter: there $\text{Cov}(x_{i,t+h}, \epsilon_{it})/\sigma^2_\vartheta = -\omega\eta^{h+1}_\lambda$, because a firm that has neither been hit nor freely reset since the shock carries all of it in its gap; here a firm hit at t carries, on average across the hit population, only the part ι^{GL} of the innovation that its non-adjusters absorb.

The weighting representation (B.51) of Section B.5.3 explains this result. Take ϵ^z_{it} as the regressor, so that it plays the role of ϵ_{it} and $\mathcal{D}^z$ collects every other draw: with probability η it equals zero (no hit), and with probability $1 - \eta$ it is uniform on $[-\varsigma/(2\omega), \varsigma/(2\omega)]$ and triggers reconsideration; conditional on a hit the firm adjusts with probability $1 - 2\bar{x}/\varsigma$ and its expected price equals the reset price (Lemma B.2). The average derivative is A^z_h, the derivative of the cross-sectional average price with respect to a shift δ of every firm's ϵ^z_{it}, i.e. a shift $\omega\delta$ of every desired price in (B.38), computed in expectation over $\mathcal{D}^z$ through Lemma B.2. A firm hit at t, or one with a free draw, passes $\omega\delta$ through fully at every horizon in that expectation. A firm that is neither is displaced by $\omega\delta$ inside its band, does not reconsider its price at t (assumption (A4)), and passes $\omega\delta$ through once it is hit or free, which by $t + h$ has probability $1 - \eta^h_\lambda$. Hence

$$\text{(B.60)} \qquad A^z_h = \omega[(1 - \eta_\lambda)\cdot 1 + \eta_\lambda\,(1 - \eta^h_\lambda)] = \omega(1 - \eta^{h+1}_\lambda) = [J^{\text{GL}}_{P,mc}\mathbf{1}]_h,$$

the pass-through of mc, as step (a) requires. The regression instead computes (B.51) with

the weight ϵ^z_{it}: the atom at zero receives no weight, so the non-hit firms, which carry the mass η_λ of zero impact responses in (B.60) (the free adjusters among them respond, but are discarded too), drop out, and the coefficient is the pass-through among hit firms, $b^z_h = \omega(1-\iota^{\mathrm{GL}}\eta^h_\lambda)$ derived above. The discrepancy $b^z_h - [J^{\mathrm{GL}}_{P,mc}\mathbf{1}]_h = \omega\eta^h_\lambda(\eta_\lambda - \iota^{\mathrm{GL}})$ is, up to the small selection term ι^{GL}, ω times the mass η^{h+1}_λ of firms neither hit nor freely reset through $t+h$; at $h=0$ it is, net of ι^{GL}, ω times the discarded mass η_λ. The identifying shock must therefore be absorbed the way mc is: it must enter p^* with the same weight ω as mc and leave the adjustment decision to other forces, so that adjusters and non-adjusters enter the regression in their population proportions. The small shifter ϑ does this in (GL) for any distribution because it never triggers adjustment. In the language of Lemma B.4, the turbulence regressor violates both conditions: it has an atom and is not Gaussian, so (G) fails, and the $\mathcal{D}$-averaged response is not affine in it, so (L) fails.

B.6. Non-Random-Walk Shocks

The main text takes z_{it} and ϑ_{it} to be driftless random walks in every model. This section investigates the case of an AR(1) shifter, $\vartheta_{it} = \rho\vartheta_{i,t-1} + \epsilon_{it}$ with $0 \le \rho < 1$. A persistent component of the z_{it} shock, with persistence ρ_z, is treated in parallel where it matters. With the price level fixed,

$$\mathbb{E}_t[p^*_{i,t+s}] - p^*_{it} = -\omega\,(1-\rho^s)\,\vartheta_{it}, \tag{B.61}$$

so two things change in general: the reset price is no longer the current desired price, and the conditional law of future desired-price increments depends on the current level ϑ_{it}. On the econometric side the regression is unchanged, and Remark 1 gives the formulas that replace (25)–(26): $\phi^{\mathrm{IRF}} = b_1/b_0 - \rho$ and $\kappa^{\mathrm{IRF}} = b_0(1-\beta\phi^{\mathrm{IRF}}\rho)/\phi^{\mathrm{IRF}}$, with ρ estimated from the autocorrelation of the observed shifter.

A benchmark for time-dependent models. The following lemma states what the geometric extrapolation recovers under any time-dependent technology, and hence what the canonical menu cost policy would predict.

Lemma B.5 (TD response to a persistent shifter). *In a time-dependent model with survival*

function Φ and the reset rule underlying (32)*, the response of prices to the shifter innovation is*

$$b_h = \omega\, c(\rho)\, \frac{\sum_{j=0}^{h} \rho^{\,h-j}\, \Phi_j}{\sum_{s\geq 0} \Phi_s}, \qquad c(\rho) \equiv \frac{\sum_{s\geq 0} \beta^s \Phi_s \rho^s}{\sum_{s\geq 0} \beta^s \Phi_s}, \tag{B.62}$$

so that $b_1/b_0 - \rho = \Phi_1$ exactly. For a mixture $\sum_k \varrho_k J^{\text{TD}}_{P,p^*}(\Phi^k)$, $b_1/b_0 - \rho$ is a weighted average of the Φ^k_1 with weights proportional to $\varrho_k c_k(\rho)/\sum_s \Phi^k_s$, which reduces to $\bar{\Phi}_1$ at $\rho = 1$.

Proof. A firm resetting at $t+j$ sets

$$p^{\text{new}}_{i,t+j} = \frac{\sum_{s\geq 0} \beta^s \Phi_s\, \mathbb{E}_{t+j} p^*_{i,t+j+s}}{\sum_{s\geq 0} \beta^s \Phi_s},$$

and $\partial \mathbb{E}_{t+j} p^*_{i,t+j+s}/\partial \epsilon_{it} = \omega \rho^{j+s}$, so $\partial p^{\text{new}}_{i,t+j}/\partial \epsilon_{it} = \omega \rho^j c(\rho)$. The expected price at $t+h$ averages reset prices over the age distribution of prices, with weight $\Phi_j/\sum_s \Phi_s$ on the reset price set j periods earlier, and only resets at dates $t+h-j \geq t$ respond; differentiating gives (B.62), and $b_1/b_0 = (\rho\Phi_0 + \Phi_1)/\Phi_0$ with $\Phi_0 = 1$. For a mixture, b_h is the ϱ_k-weighted sum of (B.62) across components, and the ratio follows. □

(C) Calvo. Nothing changes in the structure: adjustment is exogenous and the reset price is a linear forecast of desired prices,

$$\bar{\mathbf{p}}_i = J^{\text{C}}_{P,p^*}(\alpha)[\omega(\mathbf{mc} + \mathbb{E}\mathbf{z}_i + \mathbb{E}\boldsymbol{\vartheta}_i) + (1-\omega)\mathbf{P}],$$

in which ϑ_{it} enters with the weight $w_\rho \equiv (1-\beta\alpha)/(1-\beta\alpha\rho)$ and the law of z is irrelevant. Lemma B.5 with $\Phi_j = \alpha^j$ gives $c(\rho) = w_\rho$ and

$$b_h = \omega\, w_\rho\, (1-\alpha) \sum_{j=0}^{h} \rho^{\,h-j} \alpha^j, \qquad b_0 = \omega(1-\alpha) w_\rho, \qquad \frac{b_1}{b_0} = \alpha + \rho, \tag{B.63}$$

so $\phi^{\text{IRF,C}} = \alpha$ and $\kappa^{\text{IRF,C}} = b_0(1-\beta\alpha\rho)/\alpha = \omega\kappa(\alpha)$, the formulas of Remark 1. They hold exactly and for any law of ϵ_{it}, since the draw-averaged response is affine (case (L) of Lemma B.4). The same holds for any process for ϑ: $\tilde{\boldsymbol{\vartheta}}$ is then its impulse response and $b_h = \omega[J^{\text{C}}_{P,p^*}(\alpha)\tilde{\boldsymbol{\vartheta}}]_h$.

(GL) Gertler–Leahy. Neither component needs to be a random walk. Let turbulence mean-revert between hits, $z_{it} = \rho_z z_{i,t-1}$ with the uniform jump added at a hit, and let the shifter be the AR(1) above. A firm that reconsiders its price at t still chooses p to minimize $\mathbb{E}_t \sum_{s\geq 0}(\eta_\lambda\beta)^s(p - p^*_{i,t+s})^2$ along the path with no further reconsideration, so the reset price is the same linear forecast with the persistence-dependent weights

$$p_{it}^{\text{new}} = \omega\, w_{\rho_z}^{\text{GL}}\, z_{it} + \omega\, w_{\rho}^{\text{GL}}\, \vartheta_{it} + (1-\eta_\lambda\beta)\,\mathbb{E}_t \sum_{s\geq 0}(\eta_\lambda\beta)^s[\omega\, mc_{t+s} + (1-\omega)P_{t+s}], \tag{B.64}$$

with $w_r^{\text{GL}} \equiv (1-\eta_\lambda\beta)/(1-\eta_\lambda\beta\, r)$. The adjustment decision still depends on the would-be gap $y = p_{i,t-1} - p_{it}^{\text{new}}$ alone. Write $d_s \equiv p_{it}^{\text{new}} - \mathbb{E}_t p^*_{i,t+s}$ for the expected drift of the desired price away from the reset price, nonzero now. The value of keeping the old price is $-\frac{1}{2}\sum_s(\eta_\lambda\beta)^s\mathbb{E}_t(y + d_s + u_s)^2$ with u_s the unexpected part, and expanding the square, the cross term $2y\sum_s(\eta_\lambda\beta)^s d_s$ vanishes by the first-order condition that defines p_{it}^{new}; what is left is $-\frac{1}{2}y^2/(1-\eta_\lambda\beta)$ plus terms that do not involve y and are common to the keep and reset branches. Value matching therefore gives the same band (A.8), and the levels of z_{it} and ϑ_{it} enter the decision only through p_{it}^{new}: the firm's state remains one-dimensional because it does not reconsider between hits and free draws, and because the loss is quadratic. The rest of Appendix B.3 goes through with two changes. In the selection lemma, the drift of $p^{\text{new}} - \omega z$ between two hits now includes the expected reversion of the shifter, which (A2) bounds, and the turbulence draw enters the would-be gap with the weight $\omega w_{\rho_z}^{\text{GL}}$, so the uniform support in (A3) is scaled by $w_{\rho_z}^{\text{GL}}$ (with $\rho_z = 1$ nothing changes). In the aggregation step, $\partial\bar{p}_{i,t+j}^{\text{new}}/\partial\epsilon_{it} = \omega w_\rho^{\text{GL}}\rho^j$, and Lemma B.5 with $\Phi_j = \eta_\lambda^j$ gives

$$b_h = \omega\, w_\rho^{\text{GL}}\,(1-\eta_\lambda)\sum_{j=0}^{h}\rho^{\,h-j}\eta_\lambda^j, \qquad \phi^{\text{IRF,GL}} = \frac{b_1}{b_0} - \rho = \eta_\lambda, \qquad \kappa^{\text{IRF,GL}} = \omega\kappa(\eta_\lambda), \tag{B.65}$$

to the order of the approximation.

(MC) Canonical menu cost. The firm reconsiders its price every period and the value of waiting depends on where the desired price is expected to go relative to the band. The

state is the pair (y, ϑ) and the Bellman equation (B.11) becomes

$$
\begin{aligned}
V(y,\vartheta) &= \lambda \min_x W(x,\vartheta) + (1-\lambda)\min\Big\{W(y,\vartheta),\ \xi + \min_x W(x,\vartheta)\Big\},\\
W(x,\vartheta) &= \frac{1}{2}x^2 + \beta\,\mathbb{E}[V(x-v',\rho\vartheta+\epsilon') \,\big|\, \vartheta], \qquad v' = \omega(\epsilon^{z\prime} + \epsilon' - (1-\rho)\vartheta),
\end{aligned}
\tag{B.66}
$$

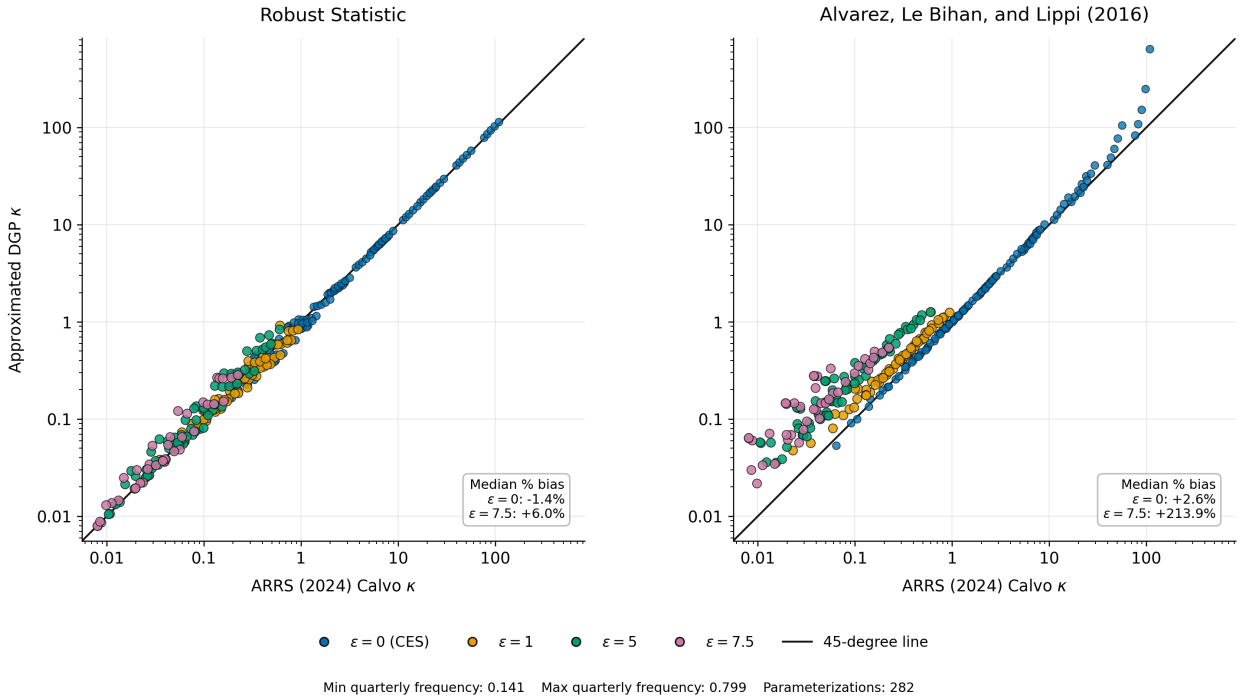
where the desired-price innovation v' has conditional mean $-\omega(1-\rho)\vartheta$. The reset gap $x^*(\vartheta)$ and the band $[\underline{x}(\vartheta), \bar{x}(\vartheta)]$ depend on ϑ: a firm with a high shifter expects its desired price to fall, so it resets below the current desired price, $x^*(\vartheta) < 0$, and tolerates larger negative than positive gaps, since a negative gap is expected to close on its own. The stationary law is a joint law of (y, ϑ).

Proposition 1 still applies and implies $b_h = [J^m_{P,mc}\tilde{\boldsymbol{\vartheta}}]_h$, $\tilde{\boldsymbol{\vartheta}} = (1, \rho, \rho^2, \dots)'$. However, the analytical results on the entries of $J^m_{P,mc}$ in Section 3.3.3 rest on the one-dimensional state and do not apply.

Appendix C. Additional Numerical Results

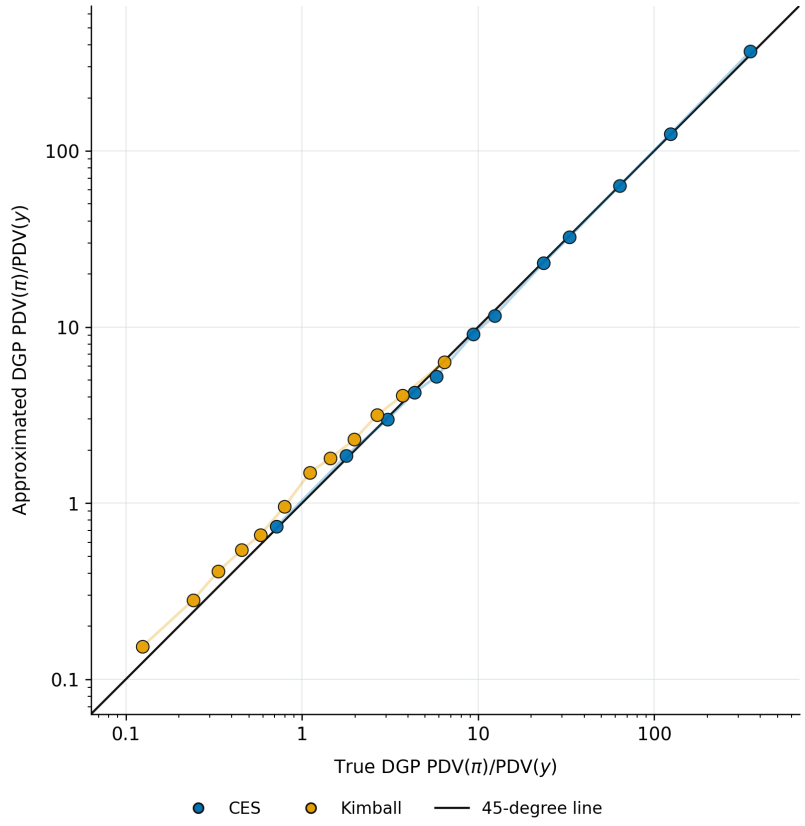
In this section we present additional numerical results.

FIGURE C.1. First Entry of the Generalized Phillips Curve with Strategic Complementarities



Notes: Each point represents one parameterization. The horizontal axis reports the Phillips-curve slope inferred from the ARRS best Calvo approximation to the menu-cost Jacobian $J^{\mathrm{C}}_{\pi,mc^r}(\phi^{\mathrm{ARRS}})$. The vertical axis reports the slope inferred from the Robust Statistic $J^{\mathrm{C}}_{\pi,mc^r}(\phi^{\mathrm{IRF}})$ (left) or the Alvarez, Le Bihan, and Lippi (2016) statistic $J^{\mathrm{C}}_{\pi,mc^r}(\phi^{\mathrm{ABL}})$ (right). Colors denote superelasticity ε, with $\varepsilon = 0$ corresponding to CES. The black line is the 45-degree line. Annotations report the median of $100(\kappa^a/\kappa^{\mathrm{ARRS}} - 1)$ for $\varepsilon = 0$ and $\varepsilon = 7.5$, where a denotes the approximation method. The sample includes 113 CES and 169 Kimball parameterizations; retained parameter coverage differs across superelasticities.

FIGURE C.2. Phillips Multiplier Under CES and Kimball



Min quarterly frequency: 0.141 Max quarterly frequency: 0.799
Parameterizations: 282

Notes: The horizontal axis reports $\text{PDV}(\pi)/\text{PDV}(y)$ in the menu-cost model. The vertical axis reports the corresponding ratio obtained by applying the generalized Phillips curve implied by the Robust Statistic $J^{\text{IRF}}_{\pi,mc^r}$ to the true model's real marginal-cost path. Both ratios use the same output path. Separately for CES and Kimball, parameterizations are sorted by the true ratio and divided into 12 approximately equal-count bins. Each point reports the arithmetic means of the true and approximated ratios within its bin. Both axes are logarithmic.