New Pricing Models, Same Old Phillips Curves?

Adrien Auclert* Rodolfo Rigato† Matthew Rognlie‡ Ludwig Straub§

May 2023

Abstract

We show that, in a broad class of menu cost models, the first-order dynamics of aggregate inflation in response to arbitrary shocks to aggregate costs are nearly the same as in Calvo models with suitably chosen Calvo adjustment frequencies. We first prove that the canonical menu cost model is first-order equivalent to a mixture of two time-dependent models, which reflect the extensive and intensive margins of price adjustment. We then show numerically that, in any plausible parameterization, this mixture is well-approximated by a single Calvo model. This close numerical fit carries over to other standard specifications of menu cost models. Thus, for shocks that are not too large, the Phillips curve for a menu cost model looks like the New Keynesian Phillips curve, but with a higher slope.

*Stanford University, CEPR and NBER. Email: aauclert@stanford.edu.
†Harvard University. Email: dinisrigato@g.harvard.edu.
‡Northwestern University and NBER. Email: matthew.rognlie@northwestern.edu.
§Harvard University, CEPR and NBER. Email: ludwigstraub@fas.harvard.edu.

This paper benefited from very useful comments from the editor, four anonymous referees, as well as Mark Aguiar, Fernando Alvarez, Marios Angeletos, Luigi Bocola, Marco Bonomo, Ricardo Caballero, Luca Dedola, Yoon Jo, Pete Klenow, Oleksiy Kryvtsov, John Leahy, Francesco Lippi, Virgiliu Midrigan, Raphael Schoenle, Joe Vavra, Gianluca Violante, Iván Werning, and Jesse Wursten. This research is supported by the National Science Foundation grant numbers SES-1851717 and SES-2042691 and the Domenic and Molly Ferrante award.

1 Introduction

The nature of nominal rigidities is a central question in monetary economics. In sticky-price models, monetary policy has aggregate effects because producers' prices do not immediately respond to changes in their costs. Motivated by the complex patterns of price changes observed in the micro data, macroeconomists have been building models with increasing degrees of realism, with the aim of improving our quantitative understanding of the behavior of aggregate inflation and the response of economic activity to changes in monetary policy.

Broadly speaking, existing price-setting models fall into two categories. The first category consists of tractable models in which firms have random opportunities to adjust their prices. In these "old" *time-dependent* (TD) models, the probability that a price can adjust is an exogenous function of the time elapsed since it last adjusted. The leading TD model is the Calvo model, where this probability is constant (Calvo 1983, Yun 1996).[1] In this model, the first-order dynamic relationship between inflation π_t and aggregate real marginal costs $\widehat{mc}_t$ is given by the well-known New Keynesian Phillips curve:

$$\pi_t = \kappa \cdot \widehat{mc}_t + \beta \mathbb{E}_t \left[\pi_{t+1} \right] \tag{NK-PC}$$

where $0 < \beta < 1$ is a discount factor and $\kappa > 0$ is the slope coefficient, with a higher slope indicating more flexible prices (see e.g. Woodford 2003b, Galí 2008). This single equation summarizes the aggregate implications of the Calvo price-setting model for small shocks to marginal costs. Its simplicity and tractability have made it ubiquitous in the New Keynesian DSGE literature.

In the past two decades, the increasing availability of administrative micro data underlying national price indices—as first documented in Bils and Klenow (2004) and Nakamura and Steinsson (2008)—has laid bare the deficiencies of TD models vis-à-vis the data, and spurred the development of a second category of price-setting models.[2] These models assume the presence of heterogeneous producers that are subject to idiosyncratic productivity shocks and adjust their prices in a lumpy fashion because of fixed "menu" costs and other features (e.g. Golosov and Lucas 2007, Klenow and Kryvtsov 2008, Nakamura and Steinsson 2010, and Midrigan 2011).[3] In these "new" *state-dependent* (SD) models, price changes are endogenous and depend both on the state of the economy and the state of the firm. Except in certain special cases, SD models must be solved numerically, and their computational complexity makes them difficult to incorporate into broader DSGE models. In particular, macroeconomists' understanding of their aggregate implications has been limited by the lack of an available equivalent to the New Keynesian Phillips curve.[4]

[1] Another widely used example of a time-dependent model is Taylor (1979). Whelan (2007), Sheedy (2010), and Carvalho and Schwartzman (2015) study these models in more generality.

[2] In the data, there is little connection between the size of price changes and the duration of price spells; and the frequency of price changes tends to move with the aggregate inflation rate. Both of these facts are inconsistent with a Calvo model but consistent with menu cost models; see for instance Klenow and Malin (2010).

[3] In addition to menu costs, these models consider random free adjustments, infrequent and leptokurtic shocks, multiple sectors, and/or multi-product firms. In turn, they build on an earlier theoretical literature that studied menu costs in partial equilibrium (e.g. Barro 1972 and Sheshinski and Weiss 1977), or general equilibrium (e.g. Caplin and Leahy 1991).

[4] Instead, to study aggregate implications of menu cost models, the literature has had to resort to stark general

In this paper, we fill this gap. We extend the notion of an aggregate, linearized Phillips curve to menu cost models and establish two major new results. First, the Phillips curve of the canonical menu cost model is the same as that of a mixture of two TD models—an *exact first-order equivalence* between SD and TD. Second, for a wide range of common parameterizations, this linearized Phillips curve is numerically almost identical to the Calvo Phillips curve (NK-PC), for some κ. This *numerical first-order equivalence* between SD and Calvo extends to broader menu cost models beyond the canonical model.

Our starting point is a formalization of the concept of a Phillips curve for a general price-setting model, as a mapping between impulse responses. Given an impulse response $(\widehat{mc}_0, \widehat{mc}_1, \ldots)$ of aggregate real marginal costs, any price-setting model produces an impulse response of inflation $(\pi_0, \pi_1, \ldots)$. We call *generalized Phillips curve* the derivative of the mapping between the two, evaluated at the steady state. This derivative, or sequence-space Jacobian (Auclert, Bardóczy, Rognlie and Straub 2021), is a linear map: for instance, if the price setting model is the standard Calvo model, this linear map is $\pi_t = \kappa \sum_{j=0}^{\infty} \beta^j \widehat{mc}_{t+j}$, i.e. the Calvo Phillips curve (NK-PC) solved forward. More generally, we represent the generalized Phillips curve as an infinite-dimensional matrix $\mathbf{K}$, so that $(\pi_0, \pi_1, \ldots)' = \mathbf{K} \cdot (\widehat{mc}_0, \widehat{mc}_1, \ldots)'$.

We proceed analogously for nominal marginal costs. Any pricing model also has a sequence-space Jacobian mapping the impulse response of nominal marginal costs $\left(\widehat{MC}_0, \widehat{MC}_1, \ldots\right)$ to that of the price level $(P_0, P_1, \ldots)$. We represent this mapping with a matrix, too, and call it the *pass-through matrix* Ψ. We show that there is a one-to-one relationship between the generalized Phillips curve $\mathbf{K}$ and the pass-through matrix Ψ, and hence that both are equivalent representations of the first-order aggregate predictions of a pricing model. We characterize both of these matrices explicitly for TD models, and then proceed to analyze them for menu cost models.

Our first main result characterizes $\mathbf{K}$ and Ψ for the canonical menu cost model. Following Alvarez, Le Bihan and Lippi (2016), we define this model to feature a quadratic loss function for producers, a Cobb-Douglas price index, permanent idiosyncratic productivity shocks, random opportunities for free adjustments, and no trend inflation. We prove that the pass-through matrix Ψ of the canonical model is a convex combination of the pass-through matrices of two fictitious TD models. Hence, the first-order aggregate implications of the canonical model are exactly the same as those of a mixture model in which a fraction of price-setters follow one TD rule and the remainder follow a different rule. In particular, the canonical model and the mixture TD model have the same generalized Phillips curve $\mathbf{K}$.

The two TD models in the mixture reflect the two margins of aggregate price adjustment in the menu cost model: first, adjustment along the *extensive* margin (movements in *Ss* adjustment bands) and second, adjustment along the *intensive* margin (movements in reset prices). Caballero and Engel (2007) previously decomposed the impact effect on the price level from a small, permanent shock to nominal marginal cost as the sum of an extensive and intensive margin component.

equilibrium assumptions, such as one-time permanent money shocks and a representative agent with a separable utility function that is linear in labor, and log in both consumption and money.

Our result shows how their structural decomposition extends to the entire impulse response of prices, and to arbitrary small shocks to costs. We show that both TD models are characterized by steady-state moments of the menu cost model and therefore theoretically recoverable from panel data on prices, and that they are easily obtained numerically, facilitating efficient computation. We also show that the adjustment hazards of these models, which we call *virtual* hazards, both eventually converge to the same constant, but generally do so from different directions: from above for the extensive margin, and from below for the intensive margin. We further show that in the presence of trend inflation, the canonical menu cost model is first-order equivalent to a mixture of *three* fictitious TD models, with two distinct models now needed to represent the extensive margin at the upper and lower adjustment bands.

An antecedent to our exact equivalence result is Gertler and Leahy (2008), which we nest as a special case. In the menu cost model studied by Gertler and Leahy (2008), firms have either no shock to their ideal price, or a shock drawn from a uniform distribution with wide support. This assumption implies that in each period, before the realization of the shock, each firm has the same probability of adjusting its price. We find that in the equivalent mixture, the extensive and intensive margin virtual hazards are constant and equal to each other. The model is therefore exactly equivalent to Calvo.

For menu cost models with a more general distribution of shocks, the Gertler and Leahy (2008) result no longer applies exactly. Our second main result, however, shows that there is still approximate *numerical equivalence* between menu cost and Calvo models. This occurs because the virtual hazards for the two TD models, while individually not constant, move in different directions and roughly offset each other in practice. Numerical equivalence is highly robust and extends beyond the canonical menu cost model. In particular, it applies to more complex pricing models, such as two-product models as in Midrigan (2011), for which our exact equivalence result no longer directly holds. It also applies to environments with modest trend inflation (up to 5%), with the Phillips curve slope steepening at higher trend inflation rates. This last result is important because it justifies using the NKPC as a model of inflation dynamics in economies with some trend inflation, even though the Phillips curve of the Calvo model with trend inflation has a different shape (Cogley and Sbordone 2008, Coibion and Gorodnichenko 2011).[5]

Our numerical equivalence result can therefore be viewed as a broad generalization of Gertler and Leahy (2008): menu cost models in a very large class, and under most reasonable parameterizations, have a generalized Phillips curve that is almost identical to the standard NKPC for some slope parameter κ. Crucially, as in Gertler and Leahy (2008), the slope κ is distinct from—and generally higher than—that implied by a Calvo model with the same adjustment frequency. That is, our numerical equivalence result focuses on the *shape* rather than the *slope* of the Phillips curve. The slope is well-understood to be affected by selection effects in price setting (e.g., Golosov and Lucas 2007). The surprising result is that these effects simply scale up the generalized Phillips

[5] In the Calvo model, trend inflation changes the shape of the NKPC by making it more forward-looking, in addition to lowering its slope. A menu cost model behaves differently—so that, at modest inflation rates, it is actually better approximated by the traditional NKPC, derived in a Calvo model *without* trend inflation. See appendix E.7.

curve.[6]

Our result is distinct from a previous connection between menu cost and Calvo models uncovered by Alvarez, Le Bihan and Lippi (2016) and further developed in Alvarez, Lippi and Passadore (2017). These papers derive an elegant sufficient statistic for the cumulative impulse response (CIR) of the price level (relative to its long-run response) to a permanent small shock to aggregate nominal costs in a broad class of pricing models, including menu cost and Calvo models. By contrast, our result shows equivalence between menu cost and Calvo for the entire impulse response of prices to any small shock to costs.[7] Our results are complementary: the Alvarez, Le Bihan and Lippi (2016) sufficient statistic gives the Calvo frequency that equalizes the *size* of monetary non-neutrality across menu cost and Calvo models; our results then establish that, for this Calvo frequency, the two models generate numerically close impulse responses.

A limitation of the Alvarez, Le Bihan and Lippi (2016) result is that it requires special general equilibrium assumptions to draw conclusions about the output effects of monetary policy.[8] By contrast, the generalized Phillips curves of menu cost models allow us to solve for the effects of monetary policy or any other small aggregate shock, in any DSGE model, under the assumption of menu cost pricing rather than Calvo pricing. We demonstrate this result in the context of two models: a textbook three-equation New Keynesian model, and the more sophisticated Smets and Wouters (2007) model. We show that correctly solving these models to first order in aggregate shocks with small idiosyncratic risk simply amounts to replacing the NKPC with the generalized Phillips curve of the menu cost model. Implementing this in a standard calibration, we find that the changes in inflation and output by switching from Calvo to menu cost pricing are negligible. We conclude that there is no loss of generality, even in the context of DSGE models, in considering the NKPC as a model of the Phillips curve, provided κ is appropriately chosen.

While it is an extremely useful benchmark, the canonical menu cost model with free adjustment is not capable of matching the rich distributions of price changes observed in micro data (Alvarez, Lippi and Oskolkov 2022a). To remedy this issue, we extend our exact equivalence result to the case in which menu costs are randomly distributed according to an arbitrary probability distribution. We then show how, with this extended result, data on price changes suffices to compute the pass-through matrix and generalized Phillips curve. This strategy can be applied to any dataset on price changes, e.g. for any country or industry, and allows researchers to derive aggregate pricing implications without needing solve a potentially complex menu cost model.[9] We illustrate this by applying the method to Israeli supermarket data from Bonomo, Carvalho, Kryvtsov, Ribon and Rigato (2022). We again find that the resulting generalized Phillips curve is

[6] Bakhshi, Khan and Rudolf (2007) numerically compare the first-generation menu cost model by Dotsey, King and Wolman (1999), which does not have idiosyncratic shocks, to a Calvo model.

[7] See Baley and Blanco (2021) for a different extension of Alvarez, Le Bihan and Lippi (2016), characterizing the CIR of higher order moments. Our focus on the entire impulse response is shared by Alvarez and Lippi (2022), who analytically characterize the impulse response to a permanent nominal marginal cost shock in a menu cost model.

[8] Under these assumptions, listed in footnote 4, the CIR of the price level to a permanent shock to nominal costs relative to its long run value is directly related to the CIR of output to permanent money shocks.

[9] Code implementing this procedure is available at `https://github.com/shade-econ/new-old-phillips-curves`.

very close to that of a Calvo model.

Our results are important for three separate literatures. First, for the literature developing partial equilibrium menu cost models that match rich aspects of the micro data, we show that solving for the generalized Phillips curve $\mathbf{K}$ allows one to embed these models into general equilibrium, and we provide three practical ways of solving for $\mathbf{K}$: an exact equivalence result, an approximate equivalence result, and a result that infers $\mathbf{K}$ directly from the price change distribution. Second, for the literature developing DSGE models, we provide a new rationalization of the Calvo Phillips curve based on menu costs, extending Gertler and Leahy (2008) to a much more general setting.

Finally, for the literature developing price-setting models to match both micro and macro data, we provide both optimism and caution. Optimism, because we can now represent menu cost models using generalized Phillips curves, which can be taken to the macro data. Caution, because these Phillips curves are so close to the Calvo model that they suffer from the same deficiencies, such as a lack of internal inflation persistence (e.g. Fuhrer and Moore 1995) and extreme forward-looking behavior (e.g. Del Negro, Giannoni and Patterson 2023). One has to look beyond menu cost models alone[10] to resolve these puzzles: for instance, to multi-sector models with complex input-output linkages (Rubbo 2023, La'O and Tahbaz-Salehi 2022), or to deviations from full-information rational expectations (Mankiw and Reis 2002, Woodford 2003a, Nimark 2008, Maćkowiak and Wiederholt 2015, Gabaix 2020, Angeletos and Huo 2021, Afrouzi and Yang 2021).[11]

The fact that our results establish a first-order connection between menu cost and TD models should not be taken to imply equivalence between menu costs and Calvo beyond first order in aggregates. When we investigate aggregate nonlinearity and state-dependence, we find a limited quantitative role for either given shocks that generate inflation of up to 5% (see appendix D.5); larger shocks, however, do lead to more price flexibility in menu cost models relative to Calvo, in line with findings from an earlier literature (e.g. Alvarez et al. 2016, Karadi and Reiff 2019). Such nonlinearities remain an important area for research. Our results also do not speak to the welfare implications of menu cost relative to Calvo models (e.g. Burstein and Hellwig 2008, Nakamura, Steinsson, Sun and Villar 2018).

In parallel and independent work, Alvarez, Lippi and Souganidis (2022b) also study the pass-through matrix Ψ of the canonical menu cost model, focusing on the continuous-time case. Their paper uses Ψ to analytically characterize the impulse response to a permanent nominal cost shock under strategic complementarities. We study strategic complementarities in section 5.2, where we show that, remarkably, they simply scale down the generalized Phillips curve $\mathbf{K}$, just like they scale down the slope parameter κ in the Calvo NKPC. Together, our papers show the importance of the pass-through matrix for the general equilibrium analysis of menu cost models.

[10] In particular, our results show that although actual adjustment hazards are increasing in menu cost models, unlike in TD models (Sheedy 2010), this does not generate inflation persistence. This is because it is the virtual hazards, rather than the actual hazards, that matter for inflation persistence.

[11] Our results also suggest that state- and time-dependent models of investment adjustment may be more closely connected than previously appreciated.

Layout. The rest of the paper is structured as follows. Section 2 sets up our benchmark time-dependent and state-dependent models, and introduces the concepts of the pass-through matrix and the generalized Phillips curve. Section 3 proves our exact equivalence result and explores its implications. Section 4 demonstrates our numerical equivalence result. Section 5 shows formally how our pricing models can be embedded into general equilibrium. Finally, section 6 shows how we can obtain the generalized Phillips curve from micro data in a richer model with generalized hazard functions.

2 Old and New Pricing Models

We begin by setting up "old" (time-dependent, TD) and "new" (state-dependent, SD) pricing models. We write the model assuming perfect foresight with respect to aggregate shocks. We then solve for first-order impulse responses to these shocks, starting from the steady state without aggregate shocks. By first-order certainty equivalence, this delivers the impulse responses to the same shocks in a fully stochastic model.[12]

2.1 State-dependent models (menu cost models)

Our benchmark state-dependent (SD) model is a discrete-time menu cost model with random free adjustments in the spirit of Nakamura and Steinsson (2010)'s "CalvoPlus" model. Like Alvarez, Le Bihan and Lippi (2016), we consider a quadratic loss function for producers, permanent idiosyncratic productivity shocks, random opportunities for free adjustments, and no trend inflation. Unlike them, we work in discrete time and allow for arbitrary distributions for the idiosyncratic shocks. We call the resulting model the "canonical menu cost model". In section 5, we justify the quadratic approximation in the context of a fully-microfounded New Keynesian model with menu-cost pricing.

There is a continuum of firms $i \in [0,1]$, each of which sells a single product in each period $t = 0, 1, 2, \ldots$, at log price p_{it} in period t. We denote by $p_{it}^* + \log MC_t$ firm i's optimal log price in period t. MC_t is the economy-wide level of the nominal marginal cost (for instance, in simple models, this would be the aggregate nominal wage). p_{it}^* captures the influence of idiosyncratic shocks on the optimal price, which can stem from idiosyncratic productivity or demand shocks.

[12] For instance, suppose that in the first-order solution to the model where the primitive innovations to aggregates are $\{\epsilon_t\}$ and aggregate nominal marginal costs follow $\log MC_t = \sum_{j=0}^{\infty} \widehat{MC}_j \epsilon_{t-j}$, the aggregate price index follows $\log P_t = \sum_{j=0}^{\infty} \widehat{P}_j \epsilon_{t-j}$, as in a Wold decomposition. Then, the sequence $\{\widehat{P}_j\}_{j=0}^{\infty}$ is also the first-order impulse response to a perturbation $\{\widehat{MC}_j\}_{j=0}^{\infty}$ to the path of marginal costs, assuming perfect foresight and starting from the steady state with idiosyncratic risk but no aggregate risk. Formally, we solve for this latter concept, but thanks to this equivalence we can use the concepts interchangeably. (For instance, we simulate data from the stochastic model and run regressions from Galí and Gertler (1999) in section 4.2.) For more discussion of certainty equivalence, see e.g. Simon (1956), Fernández-Villaverde, Rubio-Ramírez and Schorfheide (2016), and Auclert et al. (2021). Note that certainty equivalence in our case is only with respect to aggregate shocks (for which we obtain the first-order perturbation solution) and not with respect to idiosyncratic shocks (for which we obtain the full nonlinear solution).

We assume that p^*_{it} evolves as a random walk,

$$p^*_{it} = p^*_{it-1} + \epsilon_{it} \tag{1}$$

where ϵ_{it} is iid over time and across firms, drawn from a mean-zero distribution with a pdf f that is symmetric, single-peaked, and continuously differentiable, with $f'(x) < 0$ for $x > 0$ and vice versa. These assumptions nest the standard case of a normal distribution.

In each period, the firm faces a quadratic loss function equal to $\frac{1}{2}(p_{it} - p^*_{it} - \log MC_t)^2$, and has to pay an extra fixed cost ξ_{it} to change its price. The fixed cost is random, $\xi_{it} \in \{0, \xi\}$, iid over time and across firms, with a free adjustment ($\xi_{it} = 0$) materializing with probability $\lambda \in [0, 1]$.

A common and convenient way to express a firm's pricing problem in this setting is in terms of the "price gap". Here, we define the price gap x_{it} relative to the idiosyncratic optimal price, $x_{it} \equiv p_{it} - p^*_{it}$. With this definition, firm i solves the following price-setting problem:

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2}(x_{it} - \log MC_t)^2 + \xi_{it} 1_{\{x_{it} \neq x_{it-1} - \epsilon_{it}\}} \right] \tag{2}$$

where the menu cost ξ_{it} has to be paid whenever $p_{it} \neq p_{it-1}$, that is, whenever the price gap x_{it} is chosen to differ from $p_{it-1} - p^*_{it} = x_{it-1} - \epsilon_{it}$.

We define the price level P_t using a Cobb-Douglas aggregator of prices p_{it}; given that p^*_{it} has a zero cross-sectional average, this is given by:

$$\log P_t = \int x_{it} di \tag{3}$$

Inflation is given by $\pi_t = \log P_t - \log P_{t-1}$. In section 5, we derive equations (2) and (3) explicitly as an approximation to a microfounded price-setting model with menu costs.

As we show formally in appendix B, the solution to problem (2) has a well-known "*Ss*" pattern: the optimal policy takes the form

$$x_{it} = \begin{cases} x^*_t & \text{with prob } \lambda \text{ or if } x_{it-1} - \epsilon_{it} \notin [\underline{x}_t, \overline{x}_t] \\ x_{it-1} - \epsilon_{it} & \text{otherwise} \end{cases} \tag{4}$$

Following the literature, we refer to $\underline{x}_t$ and $\overline{x}_t$ as the *lower* and *upper adjustment bands*, and x^*_t as the *reset point*. In general, the policies $(\underline{x}_t, \overline{x}_t, x^*_t)$ vary over time when the sequence of nominal marginal cost $\log MC_t$ does. Given the absence of trend inflation, in a steady state, $\log MC_t$ is constant, and can be normalized to $\log MC = 0$. Then, the three policies $(\underline{x}, \overline{x}, x^*)$ are constant, with a reset price of zero, $x^* = 0$, and symmetric *Ss* bands, $\overline{x} = -\underline{x} > 0$. Price gaps converge to a stationary distribution. We denote by $g(x)$ the stationary distribution of gaps $x_{it-1} - \epsilon_{it}$ after idiosyncratic shocks occur but *before* adjustment; this convention will be convenient in what follows.

We assume that the economy is in such a steady state at the beginning of $t = 0$, with a price

gap distribution given by $g(x)$, consistent with $\log MC = 0$. We denote the probability (frequency) of price adjustment in the steady state by freq. Since price resets occur when prices leave the adjustment bands and when free reset opportunities materialize inside the adjustment bands, we have $\text{freq} = \int_{-\infty}^{\underline{x}} g(x)dx + \int_{\bar{x}}^{\infty} g(x)dx + \lambda \int_{\underline{x}}^{\bar{x}} g(x)dx$. With this notation, we can write the equation characterizing the steady-state distribution of price gaps as:

$$g(x) = \text{freq} \cdot f(x) + (1-\lambda) \int_{\underline{x}}^{\bar{x}} f(x-\tilde{x}) g(\tilde{x}) d\tilde{x} \tag{5}$$

With probability $\text{freq} \cdot f(x)$, prices are reset to zero and then drift to x. With probability $(1-\lambda) f(x-\tilde{x})$, prices drift to x from $\tilde{x}$; there are $g(\tilde{x})$ such prices in the stationary distribution. Summing across these events delivers the mass $g(x)$ in (5). Appendices C.6 and D.5 extend this environment to consider trend inflation and discuss how our main results carry over to this setting.

2.2 Time-dependent models

For state-dependent models, price setting depends only on the firm's state; for instance, in the canonical model, this state is the price gap x_{it}. For time-dependent (TD) models, by contrast, price setting depends only on the time since last adjustment (e.g. see Whelan, 2007, Sheedy, 2010, and Carvalho and Schwartzman, 2015). Now, price setting is governed by an exogenous "survival function" Φ_s for $s = 0, 1, 2, \ldots$, which counts the fraction of firms that have not yet adjusted their price after s periods among a cohort of firms that last adjusted their price at date 0. By definition, $\Phi_0 = 1$, and $\Phi_s \in [0,1]$ is weakly decreasing in s.

Each period t, firms are randomly given opportunities to reset, based on the survival function Φ_s and the time since they last adjusted. The optimal reset price gap is then given by

$$x_t^* \equiv \arg\min_x \mathbb{E}_0 \sum_{s=0}^{\infty} \beta^s \Phi_s \frac{1}{2} \left(x - \sum_{r=1}^{s} \epsilon_{it+r} - \log MC_{t+s} \right)^2 \tag{6}$$

where $x - \sum_{r=1}^{s} \epsilon_{it+r}$ is the price gap of firm i at date $t+s$ if it starts with a price gap of x at date t and does not adjust between t and $t+s$. Observe that the argmin in (6) is common across all firms, which is why we write the reset price gap as x_t^*, independent of i. Denote by $\lambda_s \equiv (\Phi_{s-1} - \Phi_s)/\Phi_{s-1} \in [0,1]$ the adjustment hazard (or adjustment probability) at horizon $s > 0$. When $\Phi_{s-1} = 0$, we set $\lambda_s = 1$. The law of motion of individual price gaps x_{it} can then be expressed as

$$x_{it} = \begin{cases} x_t^* & \text{with probability } \lambda_s \\ x_{it-1} - \epsilon_{it} & \text{otherwise} \end{cases} \qquad \text{where } s = \text{time since last adjustment.}$$

The hazards λ_s can in principle have any shape. When they are constant, $\lambda_s = \lambda \in [0,1]$, we

obtain the standard Calvo model. Accordingly, the survival function of a Calvo model is given by $\Phi_s = (1-\lambda)^s$. Another standard TD model is the T-period Taylor (1979) model, which has a stark form of increasing hazard, with $\lambda_s = 0$ for $s < T$ and $\lambda_T = 1$.

Given x_{it}, the price index and inflation are constructed as in section 2.1. One object that will be useful below is the age distribution of prices in the steady state. Denote by a_s the share of prices that last adjusted s periods ago, that is, the share of prices with age s. This distribution satisfies $a_s = (1-\lambda_s)\, a_{s-1}$. Combining this relation with the definition of λ_s, we find that $a_s \propto \Phi_s$. Since, in addition, $\sum_{s=0}^{\infty} a_s = 1$, we see that the share of prices with age s is

$$a_s = \frac{\Phi_s}{\sum_{r=0}^{\infty} \Phi_r}$$

2.3 Aggregate dynamics: pass-through matrix

The SD and TD models defined so far have in common that, at the aggregate level, starting from the steady state distribution of price gaps, they translate a (perfect-foresight) sequence of nominal marginal costs $\{MC_t\}$ to a sequence of price levels $\{P_t\}$, through the aggregation of optimal price-setting responses of heterogeneous firms to $\{MC_t\}$. In other words, both types of models describe a mapping

$$P_t = \mathcal{P}_t\left(\{MC_s\}\right) \tag{7}$$

Implicit in the function $\mathcal{P}_t$ is a time-varying distribution of price gaps induced by the $\{MC_t\}$ sequence. We are interested in the effects of small (first-order) shocks to this sequence. Log-linearizing (7) around the steady-state with constant $MC = 1$, and denoting log deviations with a hat, we have

$$\hat{P}_t = \sum_{s=0}^{\infty} \frac{\partial \log \mathcal{P}_t}{\partial \log MC_s} \widehat{MC}_s \tag{8}$$

Here, the partial derivative $\frac{\partial \log \mathcal{P}_t}{\partial \log MC_s}$ describes the response of the price level at date t with respect to an anticipated one-time unit-size shock to marginal cost at some potentially different date s. We collect all the partial derivatives in a single matrix, which we call the *pass-through matrix* $\Psi = (\Psi_{t,s})$, with $\Psi_{t,s} \equiv \frac{\partial \log \mathcal{P}_t}{\partial \log MC_s}$. This matrix is a sequence-space Jacobian (Auclert et al. 2021, Auclert, Rognlie and Straub 2023). Stacking (8) across t into a vector-valued equation, we obtain

$$\hat{\mathbf{P}} = \Psi \cdot \widehat{\mathbf{MC}} \tag{9}$$

where $\hat{\mathbf{P}} \equiv \left(\hat{P}_0, \hat{P}_1, \hat{P}_2, \ldots\right)'$ and $\widehat{\mathbf{MC}} \equiv \left(\widehat{MC}_0, \widehat{MC}_1, \ldots\right)'$.

The pass-through matrix Ψ is the first-order representation of any SD or TD pricing model. Once Ψ is computed, (9) can be used to evaluate the impulse response of the price level $\hat{\mathbf{P}}$ with respect to an arbitrary small nominal marginal cost shock $\widehat{\mathbf{MC}}$, or alternatively, to map the coefficients of an MA for marginal costs to those of an MA for prices. For instance, the s-th column of Ψ corresponds to the dynamic price level response to an anticipated one-time shock to marginal cost

at date s. By linearity, the sum across all columns of Ψ, i.e. $\sum_{s=0}^{\infty} \Psi_{t,s}$, is the price level response to a permanent shock to nominal marginal cost, as commonly analyzed in the literature (e.g. Golosov and Lucas 2007, Alvarez et al. 2016). Long-run neutrality of money implies that this response limits to 1, $\lim_{t\to\infty} \sum_{s=0}^{\infty} \Psi_{t,s} = 1$. Flexible prices correspond to the case where Ψ equals the identity matrix, so that the price level moves one for one with the marginal cost shock, irrespective of the shape of the shock.

Pass-through matrix for a TD model. For a TD model, Ψ can be evaluated analytically as follows. The reset price gap x_t^* satisfies the first-order condition of the problem in (6),

$$x_t^* = \frac{\sum_{s\geq 0} \beta^s \Phi_s \widehat{MC}_{t+s}}{\sum_{s\geq 0} \beta^s \Phi_s} \tag{10}$$

Equation (10) shows that, as in the standard Calvo model (e.g. Galí 2008), the optimal reset price gap is a weighted average of future nominal marginal cost shocks. The weights are given by a β-discounted version of the survival function Φ_s. We refer to (10) as the *policy equation*.

From equation (3) and the TD assumption we see that the price level, in turn, is a weighted average of past reset price gaps, with the age distribution as weights,

$$\hat{P}_t = \sum_{s\geq 0} a_s x_{t-s}^* = \frac{\sum_{s\geq 0} \Phi_s x_{t-s}^*}{\sum_{s\geq 0} \Phi_s} \tag{11}$$

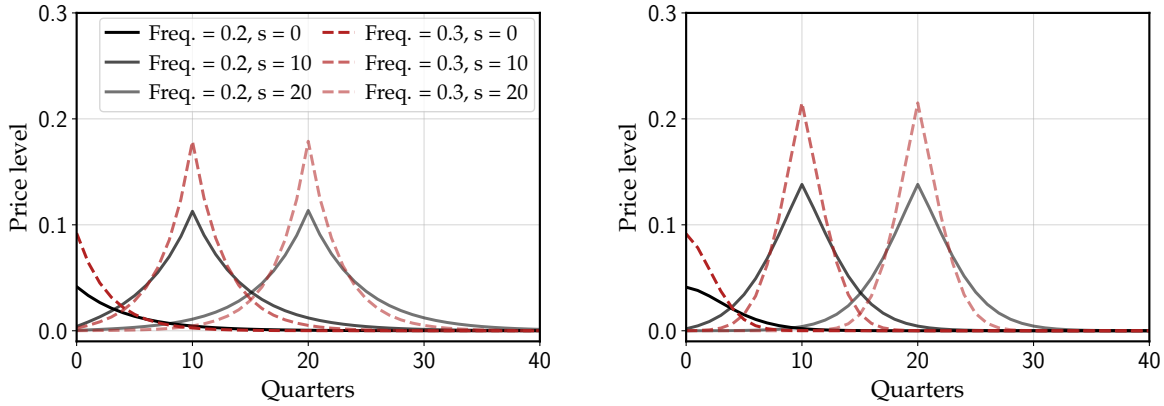
We refer to (11) as the *law of motion* of the price level. Notice that the weights in the policy equation (10) are exactly the β-discounted versions of the weights that appear in the law of motion (11). This is a key property of TD models to which we will return.

Combining the policy equation (10) and the law of motion (11), we obtain the pass-through matrix for a TD model with survival function Φ_s as:

$$\Psi^{\Phi} \equiv \frac{1}{\left(\sum_{s\geq 0} \Phi_s\right)\left(\sum_{s\geq 0} \beta^s \Phi_s\right)} \begin{pmatrix} \Phi_0 & 0 & 0 & \cdots \\ \Phi_1 & \Phi_0 & 0 & \cdots \\ \Phi_2 & \Phi_1 & \Phi_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \Phi_0 & \beta\Phi_1 & \beta^2\Phi_2 & \cdots \\ 0 & \Phi_0 & \beta\Phi_1 & \cdots \\ 0 & 0 & \Phi_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{12}$$

The matrix on the right in (12) captures the dynamic response of the reset price gap to a change in marginal costs; the matrix on the left captures the dynamic response of the price level to a change in the reset price gap.

Figure 1 displays example columns of Ψ^{Φ} in some calibrated models. The left panel shows the case of a Calvo model, where $\Phi_s = (1-\lambda)^s$ for two values of λ. The right panel shows a case of increasing adjustment hazards. The columns of Ψ^{Φ} are tent-shaped: pass-through to prices is always highest in the period of the shock itself, even if the shock is anticipated to happen at a later date $s > 0$. When the frequency of adjustment is higher, the tent is more spiked, reflecting the



(a) Calvo

(b) Increasing adjustment hazards

Figure 1: Columns $s \in \{0, 10, 20\}$ of TD pass-through matrices

Note: on the left panel, the Calvo adjustment frequencies are $\lambda \in \{0.2, 0.3\}$. On the right panel, the sequence of adjustment hazards is proportional to $1 - e^{-0.2(t+1)}$, linearly scaled in order to generate the same total adjustment frequencies $\{0.2, 0.3\}$.

fact that firms adjust less in advance and more in the period of the shock itself. With increasing hazards, the spike is more pronounced.[13]

Using the expression for Ψ^{Φ} in (12), we can evaluate the impulse responses of a TD model to arbitrary marginal cost shocks. Two important special cases are that of a one-time, perfectly transitory, shock to marginal cost at date $t = 0$; and that of a permanent shock to marginal cost. For the one-time shock, we find that the response of the price level $\hat{P}_t$ is proportional to Φ_t; for the permanent shock $\hat{P}_t$, it is proportional to the cumulative sum of Φ_t,

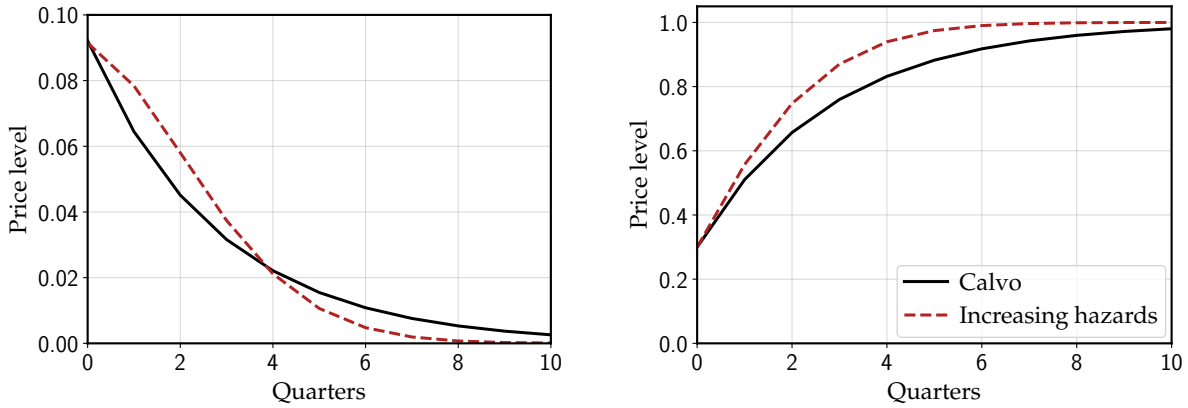
$$\text{one-time shock: } \hat{P}_t = \frac{\Phi_t}{\left(\sum_{s\geq 0} \Phi_s\right)\left(\sum_{s\geq 0} \beta^s \Phi_s\right)} \qquad \text{permanent shock: } \hat{P}_t = \frac{\sum_{s=0}^{t} \Phi_s}{\sum_{s\geq 0} \Phi_s} \tag{13}$$

Figure 2 displays these two impulse responses for the Calvo model and a model with increasing hazards. The formulas in (13) allow the survival function Φ_t of any TD model to be read off from either impulse response. In particular, the impulse response to a permanent shock delivers exactly the weights that enter the law of motion (11). We will use this property below.

2.4 Aggregate dynamics: generalized Phillips curve

The pass-through matrix characterizes the response of the price level to nominal marginal cost shocks. However, a large empirical literature (e.g. Galí and Gertler 1999, Galí, Gertler and López-Salido 2001, Sbordone 2002) studies the empirical response of inflation to *real* marginal cost, commonly known as the "Phillips curve" relationship. This distinction is also important for general equilibrium models. In simple GE models, such as that in Golosov and Lucas (2007), the pass-through matrix is sufficient to analyze the price level response to a shock to the money growth

[13] This contrast is even more visible in the Taylor (1980) model, where the spikes are triangles with height equal to the calibrated frequency (see appendix B.1).



(a) One-time shock

(b) Permanent shock

Figure 2: Responses to nominal marginal cost shocks

Note: impulse responses to one-time (left) and permanent (right) increases in nominal marginal costs. Parameter values are the same as in figure 1 for the 0.3 frequency case.

rate. But in richer models, with less restrictive assumptions on preferences and monetary shocks, a Phillips curve relationship as in (NK-PC) between inflation and real marginal cost is more useful—since, for instance, real marginal cost is closely tied to the output gap (see section 5).

We generalize this concept of a Phillips curve to a general TD or SD model as follows. Define real marginal cost as $mc_t \equiv \frac{MC_t}{P_t}$. In log-deviations, this corresponds to $\widehat{mc}_t \equiv \widehat{MC}_t - \hat{P}_t$, and in our stacked vector notation to $\widehat{\mathbf{mc}} \equiv \widehat{\mathbf{MC}} - \hat{\mathbf{P}}$. Nominal marginal cost is then given by $\widehat{\mathbf{MC}} = \widehat{\mathbf{mc}} + \hat{\mathbf{P}}$, and substituting into (9), we arrive at a fixed-point equation that relates real marginal cost and the price level,

$$\hat{\mathbf{P}} = \Psi\left(\widehat{\mathbf{mc}} + \hat{\mathbf{P}}\right) \tag{14}$$

Solving the fixed point in (14), we find a price level response equal to

$$\hat{\mathbf{P}} = \sum_{k\geq 1} \Psi^k \cdot \widehat{\mathbf{mc}} = \Psi\left(\mathbf{I} - \Psi\right)^{-1} \widehat{\mathbf{mc}} \tag{15}$$

The intuition behind (15) is that a change in real marginal cost has both direct and indirect effects, corresponding to the two terms on the right of (14). Holding the price level fixed, it directly affects nominal marginal cost, leading to an effect $\Psi \cdot \widehat{\mathbf{mc}}$ on the price level $\hat{\mathbf{P}}$. Holding $\widehat{\mathbf{mc}}$ fixed, this effect on the price level leads to a further change in nominal marginal cost and therefore on the price level. The fixed point of this problem is given by the infinite sum in (15).[14]

We can take first differences of (15) by left-multiplying both sides with $\mathbf{I} - \mathbf{L}$, where $\mathbf{L}$ is the

[14] We prove that the infinite sum $\sum_{k\geq 1} \Psi^k$ converges to $\Psi(I - \Psi)^{-1}$ in appendix D.3 for any arbitrary mixture of SD and TD models. Using a different, eigenvalue-based approach, Alvarez et al. (2022b) study the convergence properties of $\sum_{k\geq 1} (\theta\Psi)^k$ for various θ's, where θ indexes strategic complementarity in a model with nominal cost shocks. The case $\theta = 1$ is relevant for the generalized Phillips curve, since for a fixed shock to real cost, there is strategic complementarity in nominal price-setting.

lag matrix with entries of 1 one below the diagonal. This delivers the inflation response

$$\boldsymbol{\pi} = (\mathbf{I} - \mathbf{L})\,\hat{\mathbf{P}} = \underbrace{(\mathbf{I} - \mathbf{L})\,\Psi\,(\mathbf{I} - \Psi)^{-1}}_{\equiv \mathbf{K}}\,\widehat{\mathbf{mc}} \tag{16}$$

Equation (16) defines the *generalized Phillips curve* $\mathbf{K}$, or GPC for short.[15] This matrix is the linear map from an arbitrary impulse to *real* marginal cost $\widehat{\mathbf{mc}}$ to inflation. In that sense, $\mathbf{K}$ generalizes the NKPC to pricing models with a general pass-through matrix Ψ. In fact, (16) describes a one-to-one mapping between the pass-through matrix Ψ and the generalized Phillips curve $\mathbf{K}$, and the two are equivalent ways of describing the first-order behavior of prices.[16]

Generalized Phillips curve for a TD model. For a general TD model, $\mathbf{K}$ can be evaluated by first evaluating the TD pass-through matrix (12), and then implementing the formula in (16). In general this must be done numerically; however, in the case of a Calvo model with $\Phi_s = (1-\lambda)^s$, there is a particularly convenient analytical expression:

$$\mathbf{K} = \begin{pmatrix} \kappa & \beta\kappa & \beta^2\kappa & \cdots \\ 0 & \kappa & \beta\kappa & \cdots \\ 0 & 0 & \kappa & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where $\kappa = \frac{\lambda(1-\beta(1-\lambda))}{1-\lambda}$. This is the matrix version of the NKPC, which can also be written as $\pi_t = \kappa \sum_{s\geq 0} \beta^s \widehat{mc}_{t+s}$. Figure 3 plots the columns of $\mathbf{K}$ for a Calvo model (left) and for a model with increasing adjustment hazards (right). The lines in this figure have a similar interpretation as in figure 2: column s represents the inflation response to a one-time, anticipated, unit-size real marginal cost shock at date s. For a Calvo model, the inflation response is zero after date s, exactly equal to κ at date s, and discounted by β in the periods before s. With increasing hazards, there is inertia, with inflation responding less on impact, and remaining positive after date s. This inertia is due to a "catch-up" effect for prices that do not adjust when the shock hits, as discussed by Sheedy (2010).[17]

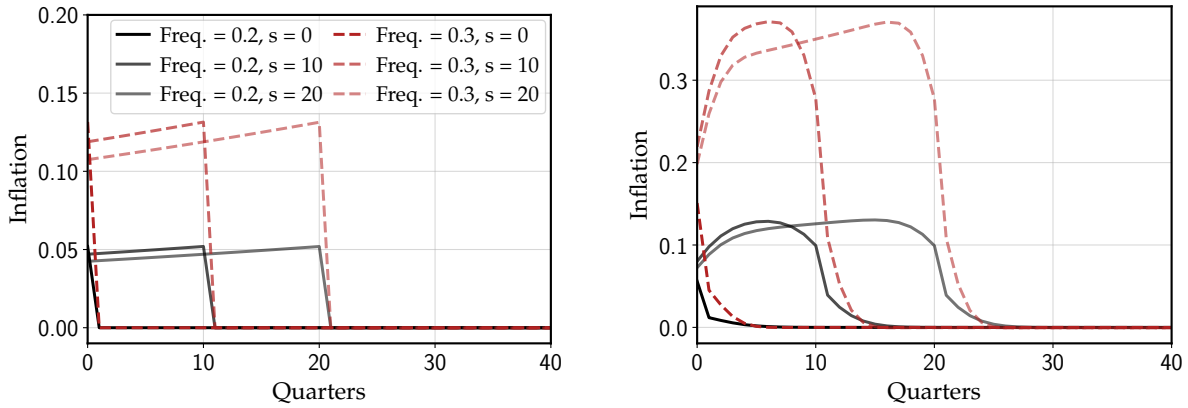
2.5 Calibration

To illustrate our theoretical results, and to provide a benchmark for our numerical results, we will simulate two SD models whose parameterizations are inspired by Golosov and Lucas (2007) (henceforth GL) and Nakamura and Steinsson (2010) (henceforth NS). We assume that the shock

[15] We pick the letter $\mathbf{K}$ in order to mirror the slope parameter κ in the NKPC.

[16] To obtain Ψ from $\mathbf{K}$, write $\Psi = \mathbf{UK}\,(\mathbf{I} + \mathbf{UK})^{-1}$ where $\mathbf{U} = (U_{ts})$ is the matrix with $U_{ts} = 1$ for $t \geq s$ and 0 elsewhere, which satisfies $\mathbf{U}\,(\mathbf{I} - \mathbf{L}) = \mathbf{I}$. Note that both $\hat{\mathbf{P}} = \Psi \cdot \widehat{\mathbf{MC}}$ and $\boldsymbol{\pi} = \mathbf{K} \cdot \widehat{\mathbf{mc}}$ hold regardless of the primitive aggregate shock—for instance, a monetary or TFP shock, both of which we will consider in section 5—as long as the shock does not itself change the TD or SD model.

[17] In appendix B.1, we show the generalized Phillips curve corresponding to the Taylor (1980) model.



(a) Calvo

(b) Increasing adjustment hazards

Figure 3: Columns $s \in \{0, 10, 20\}$ of time-dependent Phillips curve matrices

Note: Generalized Phillips curves for the same time-dependent models as in figure 1.

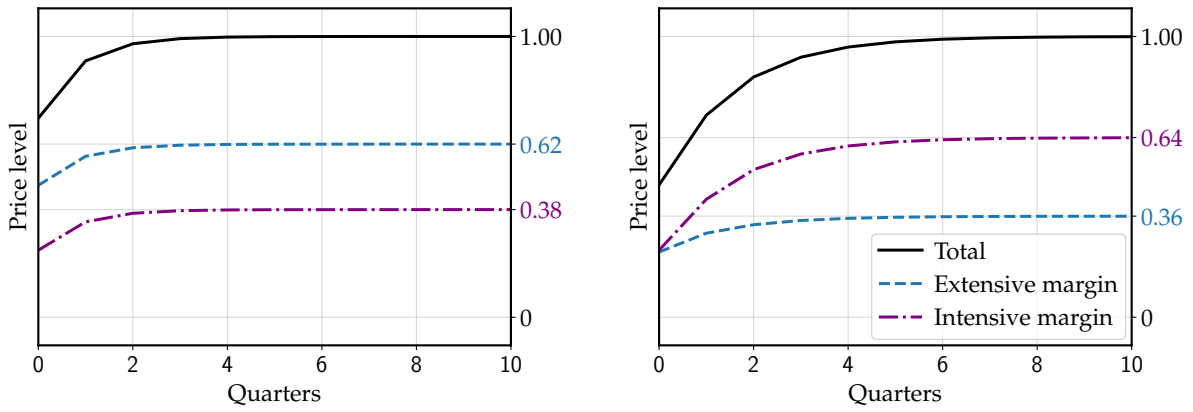
	Golosov-Lucas (GL)	**Nakamura-Steinsson** (NS)
Menu cost (ξ)	0.0060	0.0513
Prob. of free adj. (λ)	0	0.179
Shock std. (σ_ϵ)	0.046	0.060
Discount factor (β)	0.99	0.99

Table 1: Calibrated parameter values.

distribution f is normal with variance σ_ϵ^2. Up to a first-order approximation, the aggregate behavior of an SD pricing model depends only on three parameters: the menu cost normalized by the variance of idiosyncratic shocks ξ/σ_ϵ^2, the rate of free adjustments λ, and the discount factor β (see appendix B.4). In our context, this implies that the pass-through matrix Ψ and generalized Phillips curve $\mathbf{K}$ of an SD model are only a function of those three parameters. We use the following standard calibration strategy to pin down these parameters.

For GL, we assume no free adjustments, $\lambda = 0$, and choose ξ and σ_ϵ to match a quarterly average frequency of price changes of 23.9% and a median size of price adjustments of 8.5%. This corresponds to the frequency and adjustment size for the median sector in the US CPI (see Nakamura and Steinsson 2010). For NS, we keep the same targets, but also choose λ in order to match a share of free adjustments of 75%, similar to Nakamura and Steinsson (2010).[18] We set the discount factor to $\beta = 0.99$. These calibrated parameters are summarized in table 1.

[18] In the CalvoPlus model introduced by Nakamura and Steinsson (2010), firms can adjust their prices either by paying a large menu cost that is always available, or by paying a smaller menu cost that is only infrequently available. The model is calibrated so that approximately 75% of adjustments occur in the low menu cost state. Our model is, therefore, an approximation of theirs in which the smaller menu cost is set to zero.



(a) Golosov-Lucas (b) Nakamura-Steinsson

Figure 4: Price level responses to permanent nominal marginal cost shocks

Note: impulse responses of prices to a permanent nominal marginal cost shocks decomposed into changes in reset points (intensive margin) and adjustment bounds (extensive margin). Parameter values as in table 1.

3 Exact Equivalence between SD and TD Pricing Models

We are now ready to compare the aggregate implications of state- and time-dependent models for the dynamics of prices and inflation. Since the pass-through matrix encapsulates all the first-order implications of the pricing models introduced so far, a simple place to start is as follows. Consider an SD model with a given pass-through matrix Ψ. Can we find a survival function Φ_s such that the pass-through matrix Ψ^{Φ} of the TD model with survival function Φ_s is equal to Ψ?

If such a Φ_s exists, one should be able to recover it from a single impulse response in the SD model—in particular, the impulse response to a permanent shock to nominal marginal cost. The black lines in figure 4 show the impulse responses of the price level to such a shock in the GL and NS models. As expected from the literature, the GL model shows a faster convergence of the price level to 1 than the NS model.

The equation for the impulse response to a permanent shock in (13) implies that, if this SD impulse response is generated by a TD model with survival function Φ_s, then it should be equal to $\sum_{s=0}^{t}\Phi_s/\sum_{s\geq 0}\Phi_s$ at each t. This gives us a way to read off a candidate Φ_s from figure 4. Unfortunately, while this procedure will by construction generate the correct impulse response to permanent shocks, it generally does not produce the correct impulse responses to any other shock in the SD model. In other words, no single TD model can match the entire pass-through matrix of an SD model: in this sense, TD and SD models are truly different.

3.1 Exact equivalence result

To make progress, we now consider the underlying drivers of the SD impulse responses in figure 4. The permanent shock shifts up both the adjustment bands $\underline{x}_t, \overline{x}_t$, and the reset price gap x_t^*, both by the amount of the shock. As in Caballero and Engel (2007), we can separate the overall

impulse response of prices into an *extensive margin* component, driven by the shift $(d\underline{x}_t, d\bar{x}_t)$ in the adjustment bands, and an *intensive margin* component, driven by the shift dx_t^* in the reset price gap. Since the shock is evaluated to first order, this decomposition is additive.

The blue and purple lines in figure 4 show those extensive and intensive margin contributions for the GL and NS models. On impact, as Caballero and Engel (2007) showed, the contribution of the intensive margin is the same across the two models, and equal to the calibrated frequency of price changes (freq). As they also pointed out, the impact contribution of the extensive margin is large in the GL model, contributing to faster aggregate adjustment than in the NS model. The figure shows that the extensive margin continues to make a large contribution through the impulse response of the GL model, and ultimately accounts for 62% of its eventual price level response. For NS, by contrast, the long-run share of the extensive margin is only 36%.

Now consider fitting a time-dependent model as above, but separately for the extensive and intensive margin impulse responses. For each, we read off a survival function, which we denote by Φ^e and Φ^i for the extensive and intensive margins. We call these *virtual survival functions*, since they are different from the actual survival function of the SD model, as we discuss further in section 3.3 below. We can also read off the share α of the eventual price level response that is accounted for by the extensive margin.

Our main result in this section is that these survival functions are structural: the implied mixture TD model has the same first-order impulse responses to all shocks as the underlying SD model.

Proposition 1. *The pass-through matrix Ψ of the canonical menu cost model is a mixture of two time-dependent (TD) pass-through matrices,*

$$\Psi = \alpha \Psi^{\Phi^e} + (1-\alpha)\, \Psi^{\Phi^i} \tag{17}$$

The "virtual" survival functions Φ^e and Φ^i and the share α are such that:

- $\alpha \Sigma_{s=0}^{t} \Phi_s^e / \sum_{s\geq 0} \Phi_s^e$ *is the impulse response of the price level to a permanent nominal marginal cost shock when only the* Ss *band shifts;*
- $(1-\alpha)\, \Sigma_{s=0}^{t} \Phi_s^i / \sum_{s\geq 0} \Phi_s^i$ *is the impulse response of the price level to a permanent nominal marginal cost shock when only the reset price gap shifts.*

Proposition 1 formalizes this TD-matching approach. In the proof, we also provide explicit expressions for Φ^e, Φ^i, and α that can be obtained from primitives. The pass-through matrix of an SD model is exactly equal to the linear combination of the pass-through matrix of the extensive margin (with share α) and the pass-through matrix of the intensive margin (with share $1-\alpha$).[19] This implies that the SD impulse response to *any* shock, $\Psi \cdot \widehat{\mathbf{MC}}$, can be exactly decomposed into an extensive margin contribution $\alpha \Psi^{\Phi^e} \cdot \widehat{\mathbf{MC}}$ and an intensive margin contribution $(1-\alpha)\, \Psi^{\Phi^i} \cdot \widehat{\mathbf{MC}}$,

[19] Observe that the survival and hazard functions of the mixture model are not necessarily convex combinations of the individual survival and hazard functions.

and that both contributions come from TD models. For example, given equation (13), the impulse response of the SD model to a one-time shock is simply given by

$$\alpha \frac{\Phi_0^e}{\left(\sum_{s\geq 0}\Phi_s^e\right)\left(\sum_{s\geq 0}\beta^s\Phi_s^e\right)}\Phi_t^e + (1-\alpha)\frac{\Phi_0^i}{\left(\sum_{s\geq 0}\Phi_s^i\right)\left(\sum_{s\geq 0}\beta^s\Phi_s^i\right)}\Phi_t^i \tag{18}$$

Our result thus naturally generalizes the Caballero and Engel (2007) decomposition to arbitrary shocks and to the entire impulse response.[20] Since the pass-through matrix characterizes the entire first-order behavior of a pricing model, proposition 1 implies more broadly that the SD model is identical to a mixture TD model, in which a fixed share of firms α follows the TD rule Φ^e and the remaining firms follow the TD rule Φ^i.

Corollary 1. *To first order, the aggregate pricing behavior of the canonical menu cost model is identical to that of a mixture of a time-dependent model with survival function Φ^e and weight α, and a time-dependent model with survival function Φ^i and weight $1-\alpha$. In particular, these two models share the same generalized Phillips curve.*

A useful way to interpret our equivalence result is as one of dimensionality reduction. To see this, truncate the matrices in (17) to be of size $T \times T$. From (12), we see that, up to a constant, a $T \times T$ truncated TD pass-through matrix Ψ^Φ only depends on $\Phi_0, \ldots, \Phi_{T-1}$. Thus, (17) reduces the number of dimensions of the model from T^2 down to $2T-1$.[21]

Computational benefits of proposition 1. The dimensionality reduction idea highlights the computational benefits of proposition 1. It is typically relatively straightforward to compute the impulse response of the price level to permanent nominal marginal cost shocks in menu cost models—making this a popular exercise for papers in the literature (e.g., Golosov and Lucas 2007, Alvarez et al. 2016). It is typically much harder to compute the impulse responses to non-permanent, e.g. AR(1), shocks, and even harder to embed menu cost models in fully specified general-equilibrium models without making restrictive assumptions on preferences and monetary shocks.

Proposition 1 suggests a simple way to solve these computational issues, as follows. First, compute the impulse response to a permanent shock assuming that only the *Ss* bands $(\underline{x}, \overline{x})$ adjust, and then assuming that only the reset price gap x^* adjusts, obtaining Φ^e, Φ^i, α as outlined above. Then compute the right hand side of (17) using the formula in (12) to obtain the pass-through matrix Ψ. This makes it possible to simulate arbitrary shocks to nominal costs by taking the matrix

[20] In empirical work, it is also common to decompose inflation into an intensive and an extensive margin, e.g. Klenow and Kryvtsov (2008) and Dedola, Kristøffersen and Züllig (2021). Typical decompositions in this literature relate the extensive margin to movements in the frequency of price changes. In our model, the overall frequency of price changes is constant to first order, with the extensive margin term reflecting opposite-sign movements in the frequencies of price increases and declines.

[21] $2T-2$ for $\Phi_1^e, \ldots, \Phi_{T-1}^e$ and $\Phi_1^i, \ldots, \Phi_{T-1}^i$ given that $\Phi_0^e = \Phi_0^i \equiv 1$; 1 for the constants multiplying the truncated matrices in (17).

product of Ψ and the shock vector. In addition, using (16), we can construct the generalized Phillips curve $\mathbf{K}$ and simulate the response to arbitrary real marginal cost shocks as well.

Continuous time. While our result is set in discrete time, a similar result holds in continuous time. We present it in appendix A.

Trend inflation. In an environment with trend inflation, the steady state no longer features symmetric *Ss* bands. In appendix C.6, we show that this implies that the pass-through matrix Ψ becomes a mixture of three TD pass-through matrices, corresponding respectively to adjustment at the lower adjustment band, the upper adjustment band, and the reset point. As trend inflation goes to zero, the first two can be consolidated into a single TD model for the extensive margin and we recover equation (17).

3.2 Proof of proposition 1

The proof of proposition 1 has several steps. First, we introduce a new object: the expected price gap $E^t(x)$, t periods in the future, for a firm with a price gap of x today. We then study how, to first order, the aggregate price index is affected by past policy changes (the law of motion), and how these policy changes are determined in response to future marginal cost shocks (the policy equation). When we combine the law of motion and policy equation to obtain the pass-through matrix, we find that this matrix is a weighted sum of two terms, representing the extensive and intensive margins. Each of these terms has exactly the same form as in the time-dependent case (12), with survival functions derived from $E^t(x)$.

Expected price gaps. Consider the canonical menu cost model in steady state, with $\log MC = 0$. Let x be the price gap of firm i at the end of period 0, after it has had a chance to adjust. For each $t \geq 0$, we define $E^t(x) \equiv \mathbb{E}_0 [x_{it} | x_{i0} = x]$ as the firm's expected price gap at the end of period t.[22] Clearly, the identity of firm i is irrelevant for this object, so $E^t(x)$ only depends on the price gap x and the horizon t. Since the model is symmetric in steady state, E^t is an odd function for all t, i.e. $E^t(-x) = -E^t(x)$, and in particular $E^t(0) = 0$.

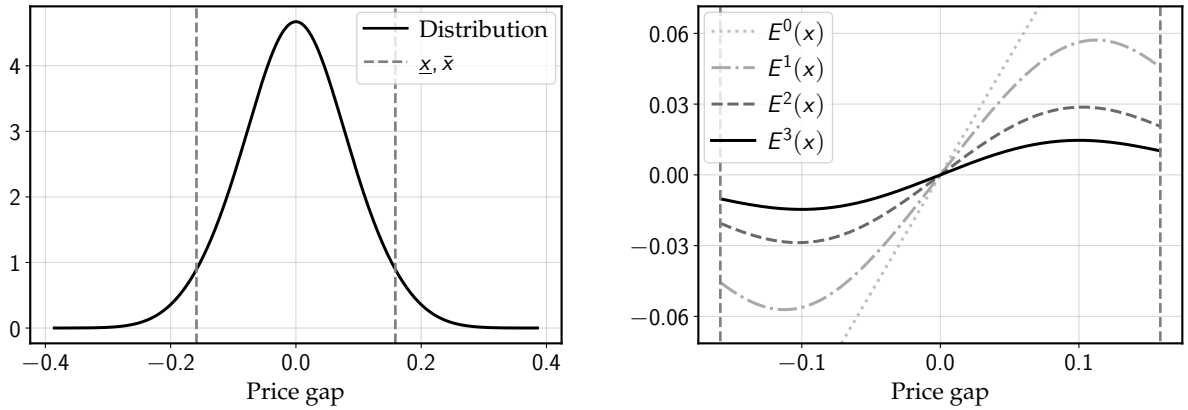
Starting with $E^0(x) = x$ and applying the law of iterated expectations, $E^t(x)$ is given recursively for $t > 0$ by

$$E^t(x) = (1-\lambda) \int_{\underline{x}}^{\bar{x}} f(x'-x) E^{t-1}(x') dx', \tag{19}$$

taking expectations over $E^{t-1}(x')$ using the no-adjustment transition probability $(1-\lambda)f(x'-x)$ from x to x'. (The contribution from price resets to (19) is 0, since $E^{t-1}(0) = 0$.)

The right panel in figure 5 plots $E^t(x)$ as function of x for different horizons t. At longer horizons t, expected price gaps all converge towards zero. This happens for two reasons: first,

[22] These objects also feature in Alvarez et al. (2016) and Alvarez and Lippi (2022), who derive an analytical expression for them in continuous time (see appendix A).



(a) Stationary distribution $g(x)$

(b) Expected price gaps $E^t(x)$

Figure 5: Stationary distribution $g(x)$ and expected price gaps $E^t(x)$

Note: stationary distribution of price gaps before adjustments (left) and expected future price gaps given initial position x (right) for the Nakamura-Steinsson model, calibrated as in table 1. The vertical dashed lines denote adjustment bands. Prices that lie outside adjustment bands, and a random fraction λ of those inside, adjust to 0.

prices are more likely to have adjusted at longer horizons, after which their expected price gaps are zero; second, the expected price gap conditional on not having adjusted also converges to zero, due to a selection effect that we explore in more detail in the next section.

The crucial object in our proof is $E^t(x)/x$. This measures the expected *persistence* of any price gap x over t periods, starting at one for $t = 0$ and ultimately converging to zero. We will show below that the virtual survival functions for the extensive and intensive margins are given by the persistences of the price gaps at the adjustment bands and reset point, respectively,

$$\Phi_t^e \equiv \frac{E^t(\overline{x})}{\overline{x}} \quad \text{and} \quad \Phi_t^i \equiv E^{t\prime}(0) \ \left(= \lim_{x \to 0} \frac{E^t(x)}{x} \right) \tag{20}$$

where the limit on the right follows from L'Hôpital's rule. The weight α on the extensive margin is given by

$$\alpha \equiv 2(1-\lambda) g(\overline{x}) \overline{x} \cdot \sum_{t \geq 0} \frac{E^t(\overline{x})}{\overline{x}} \tag{21}$$

Relationship between expected price gaps and the law of motion. In general, the log price level at any date t, after adjustment, is given by

$$\log P_t = (1-\lambda) \int_{\underline{x}_t}^{\overline{x}_t} x g_t(x) dx + \left(\lambda + (1-\lambda) \left(1 - \int_{\underline{x}_t}^{\overline{x}_t} g_t(x) dx \right) \right) x_t^* \tag{22}$$

where $g_t(x)$ is the density of price gaps at date t before adjustments in that period, and the first and second terms are the contributions from non-adjusters and adjusters, respectively.

Now suppose that at date $t - s$, starting from the stationary distribution $g(x)$, there is a one-time change in policies $\overline{x}_{t-s}$, $\underline{x}_{t-s}$, and x_{t-s}^*, after which policies all return to the steady state. Given

perfect foresight, the average price gap x at date t will be the average of expected price gaps $E^s(x)$ at date $t-s$. Using this insight, we can rewrite $\log P_t$ as

$$\log P_t = (1-\lambda)\int_{\underline{x}_{t-s}}^{\overline{x}_{t-s}} E^s(x)g(x)dx + \left(\lambda + (1-\lambda)\left(1 - \int_{\underline{x}_{t-s}}^{\overline{x}_{t-s}} g(x)dx\right)\right) E^s(x^*_{t-s}) \tag{23}$$

Relative to (22), we replaced x and x^*_t with the expected price gaps $E^s(x)$ and $E^s(x^*_{t-s})$ at date t starting from x and x^*_{t-s} at date $t-s$, respectively.

Totally differentiating (23) around the steady state, we have

$$\begin{aligned} d\log P_t = (1-\lambda)\left(E^s(\overline{x})g(\overline{x})d\overline{x}_{t-s} - E^s(\underline{x})g(\underline{x})d\underline{x}_{t-s}\right) \\ - (1-\lambda)\left(g(\overline{x})d\overline{x}_{t-s} - g(\underline{x})d\underline{x}_{t-s}\right)E^s(0) + \text{freq}\cdot(E^s)'(0)dx^*_{t-s} \end{aligned}$$

The first term simplifies due to symmetry, $E^s(\underline{x}) = -E^s(\overline{x})$ and $g(\overline{x}) = g(\underline{x})$, the second term is zero since $E^s(0) = 0$, and the third term has the stated form since $\text{freq} = \lambda + (1-\lambda)\left(1 - \int_{-\overline{x}}^{\overline{x}} g(x)dx\right)$ is the steady-state price adjustment frequency. We thus arrive at

$$d\log P_t = (1-\lambda)E^s(\overline{x})g(\overline{x})(d\overline{x}_{t-s} + d\underline{x}_{t-s}) + \text{freq}\cdot(E^s)'(0)dx^*_{t-s} \tag{24}$$

Equation (24) gives the first-order response, around the steady state, of $\log P_t$ to changes in policies at any date $t-s$. We can then sum these contributions from each $t-s \geq 0$ to obtain the full first-order law of motion for prices

$$d\log P_t = (1-\lambda)g(\overline{x})\sum_{s=0}^{t} E^s(\overline{x})(d\overline{x}_{t-s} + d\underline{x}_{t-s}) + \text{freq}\cdot\sum_{s=0}^{t}(E^s)'(0)dx^*_{t-s} \tag{25}$$

Relationship between expected price gaps and the policy equation. Let $V_t(x)$ denote the post-adjustment value function for a firm at any date t. Given equation (2), this satisfies

$$V_t(x) \equiv \frac{1}{2}(x - \log MC_t)^2 + \beta(1-\lambda)\mathbb{E}_\epsilon\left[\min(V_{t+1}(x+\epsilon), \xi + \min_{x^*} V_{t+1}(x^*))\right] + \beta\lambda\min_{x^*} V_{t+1}(x^*) \tag{26}$$

To start, suppose that aggregate marginal cost remains at its steady-state level at every date except s, where there is a shock $d\log MC_s$. Differentiating equation (26) around the steady state, this implies $dV_s(x) = -x\cdot d\log MC_s$. Further, we show in appendix C.1 that by an envelope argument, the implied perturbation to the value function $dV_t(x)$ for any $t < s$ is

$$dV_t(x) = \beta(1-\lambda)\int_{\underline{x}}^{\overline{x}} f(x'-x)dV_{t+1}(x')dx \tag{27}$$

i.e. that it is the discounted change to $dV_{t+1}(x')$, taking expectations over all x' where there is no adjustment under the steady-state policy.[23]

[23] The contribution from resets turns out to be zero because all dV_t are odd and satisfy $dV_t(0) = 0$.

Given that $dV_s(x) = -E^0(x) \cdot d\log MC_s = -x \cdot d\log MC_s$, and that (27) has exactly the same form as our earlier recursion (19), except with an additional discount factor β, It follows that

$$dV_t(x) = -\beta^{s-t} E^{s-t}(x) \cdot d\log MC_s. \tag{28}$$

Using similar arguments, we show in appendix C.1 that $V'(x) = \sum_{u=0}^{\infty} \beta^u E^u(x)$.

At each date t, the optimal adjustment thresholds are given by value-matching conditions $V_t(\bar{x}_t) = V_t(\underline{x}_t) = V_t(x_t^*) + \zeta$, and the optimal reset point is given by the first-order condition $V_t'(x_t^*) = 0$. Totally differentiating around the steady state, we have $d\bar{x}_t = -\left(dV_t(\bar{x}) - dV_t(0)\right)/V'(\bar{x})$, $d\underline{x}_t = -\left(dV_t(\underline{x}) - dV_t(0)\right)/V'(\underline{x})$, and $dx_t^* = -dV_t'(0)/V''(0)$, which combined with our results above become $d\bar{x}_t = d\underline{x}_t = \frac{\beta^{s-t}E^{s-t}(\bar{x})}{\sum_{u=0}^{\infty}\beta^u E^u(\bar{x})} d\log MC_s$ and $dx_t^* = \frac{\beta^{s-t}(E^{s-t})'(0)}{\sum_{u=0}^{\infty}\beta^u (E^u)'(0)} d\log MC_s$.

Allowing for shocks at different dates s and summing to get the overall effect, we conclude

$$d\bar{x}_t = d\underline{x}_t = \frac{\sum_{s\geq t}\beta^{s-t}E^{s-t}(\bar{x})d\log MC_s}{\sum_{s\geq t}\beta^{s-t}E^{s-t}(\bar{x})} \tag{29}$$

$$dx_t^* = \frac{\sum_{s\geq t}\beta^{s-t}(E^{s-t})'(0)d\log MC_s}{\sum_{s\geq t}\beta^{s-t}(E^{s-t})'(0)} \tag{30}$$

i.e. that both the changes in thresholds $d\bar{x}_t, d\underline{x}_t$ and changes in reset point dx_t^* are given by weighted averages of shocks to future marginal cost.

Writing as mixture of time-dependent models. In vector form, we can write our law of motion for prices (25), given that $d\bar{\mathbf{x}} = d\underline{\mathbf{x}}$, as

$$\hat{\mathbf{P}} = 2(1-\lambda)g(\bar{x}) \begin{pmatrix} E^0(\bar{x}) & 0 & 0 & \cdots \\ E^1(\bar{x}) & E^0(\bar{x}) & 0 & \cdots \\ E^2(\bar{x}) & E^1(\bar{x}) & E^0(\bar{x}) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} d\bar{\mathbf{x}} + \text{freq} \cdot \begin{pmatrix} E^{0\prime}(0) & 0 & 0 & \cdots \\ E^{1\prime}(0) & E^{0\prime}(0) & 0 & \cdots \\ E^{2\prime}(0) & E^{1\prime}(0) & E^{0\prime}(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} d\mathbf{x}^*$$

where $\overline{\mathbf{x}} \equiv (\overline{x}_0, \overline{x}_1, \overline{x}_2, \ldots)'$, etc. Then, substituting in the vector form of (29)–(30) and rearranging, this becomes

$$\hat{\mathbf{P}} = \frac{2(1-\lambda)g(\overline{x})\overline{x}\sum_{s\geq 0}\frac{E^s(\overline{x})}{\overline{x}}}{\left(\sum_{s\geq 0}\frac{E^s(\overline{x})}{\overline{x}}\right)\left(\sum_{s\geq 0}\beta^s\frac{E^s(\overline{x})}{\overline{x}}\right)}\begin{pmatrix} \frac{E^0(\overline{x})}{\overline{x}} & 0 & 0 & \cdots \\ \frac{E^1(\overline{x})}{\overline{x}} & \frac{E^0(\overline{x})}{\overline{x}} & 0 & \cdots \\ \frac{E^2(\overline{x})}{\overline{x}} & \frac{E^1(\overline{x})}{\overline{x}} & \frac{E^0(\overline{x})}{\overline{x}} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} \frac{E^0(\overline{x})}{\overline{x}} & \beta\frac{E^1(\overline{x})}{\overline{x}} & \beta^2\frac{E^1(\overline{x})}{\overline{x}} & \cdots \\ 0 & \frac{E^0(\overline{x})}{\overline{x}} & \beta\frac{E^1(\overline{x})}{\overline{x}} & \cdots \\ 0 & 0 & \frac{E^0(\overline{x})}{\overline{x}} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\widehat{\mathbf{MC}}$$
$$+\frac{\text{freq}\cdot\sum_{s\geq 0}E^{s\prime}(0)}{\left(\sum_{s\geq 0}E^{s\prime}(0)\right)\left(\sum_{s\geq 0}\beta^s E^{s\prime}(0)\right)}\begin{pmatrix} E^{0\prime}(0) & 0 & 0 & \cdots \\ E^{1\prime}(0) & E^{0\prime}(0) & 0 & \cdots \\ E^{2\prime}(0) & E^{1\prime}(0) & E^{0\prime}(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} E^{0\prime}(0) & \beta E^{1\prime}(0) & \beta^2 E^{2\prime}(0) & \cdots \\ 0 & E^{0\prime}(0) & \beta E^{1\prime}(0) & \cdots \\ 0 & 0 & E^{0\prime}(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\widehat{\mathbf{MC}} \tag{31}$$

We see that each term above, aside from the numerators, has *exactly* the form of the time-dependent pass-through matrix (12). In particular, just as in the time-dependent case, the rows of the upper triangular matrices (representing the policy equation that maps marginal cost shocks to policies) are the same as the columns of the lower triangular matrices (representing the law of motion that maps policy shocks to aggregate prices), except with added discounting.[24] Hence, if we define the survival functions as in (20), this pass-through matrix simplifies to just

$$\Psi = 2(1-\lambda)g(\overline{x})\overline{x}\left(\sum_{s\geq 0}\frac{E^s(\overline{x})}{\overline{x}}\right)\Psi^{\Phi^e} + \text{freq}\left(\sum_{s\geq 0}E^{s\prime}(0)\right)\Psi^{\Phi^i}, \tag{32}$$

a weighted sum of the time-dependent pass-through matrices Ψ^{Φ^e} and Ψ^{Φ^i}. The first term gives the response from extensive-margin adjustments to $\overline{x}_t$ and $\underline{x}_t$, and the second term gives the response from intensive-margin adjustments to x_t^*.

In response to a permanent shock to marginal cost, it follows directly from (12) that the long-term response of prices must be one-for-one in time-dependent models, and we also know that it must be one-for-one in our state-dependent model.[25] For (32) to be consistent with this, the sum of the coefficients on Ψ^{Φ^e} and Ψ^{Φ^i} must equal 1.[26] Hence, defining α as in (21), we can rewrite (32) as just

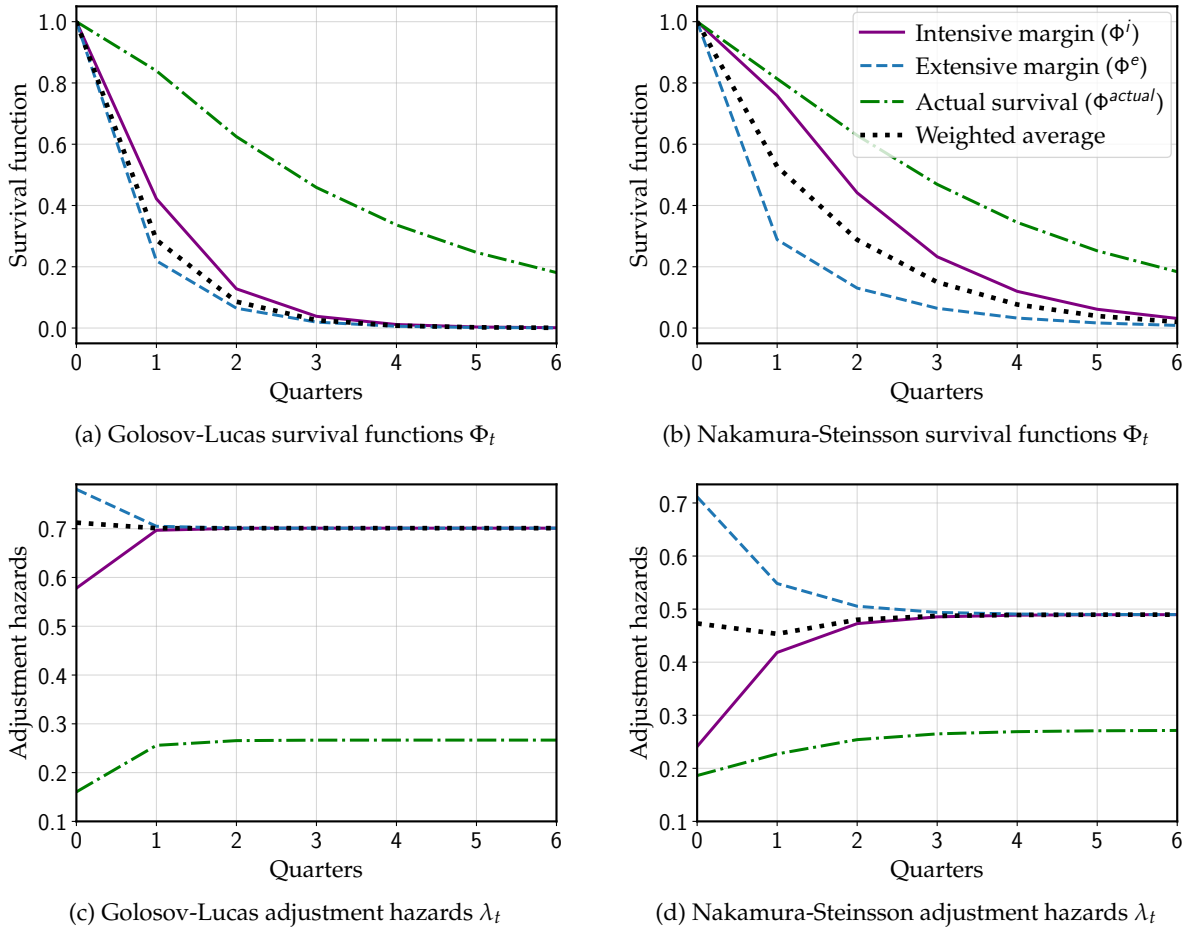
$$\Psi = \alpha\Psi^{\Phi^e} + (1-\alpha)\Psi^{\Phi^i} \tag{33}$$

which is identical to (17) in proposition 1.

[24] Also, the first entries in these sequences, $\frac{E^0(\overline{x})}{\overline{x}}$ and $E^{0\prime}(0)$, equal 1, and appendix C.2 proves that they are positive and decreasing, as required for survival functions.

[25] It follows immediately from (29)–(30) that a unit permanent increase in marginal cost results in a unit permanent increase in both adjustment thresholds and the reset point. Just as in the original steady state, price gaps eventually converge to the new ergodic distribution, identical but translated to the right by these increases.

[26] It is also possible to prove this result directly, starting from the expression in (32). We owe the proof, provided in appendix C.3, to an anonymous referee.



(a) Golosov-Lucas survival functions Φ_t

(b) Nakamura-Steinsson survival functions Φ_t

(c) Golosov-Lucas adjustment hazards λ_t

(d) Nakamura-Steinsson adjustment hazards λ_t

Figure 6: Survival functions and adjustment hazards

Note: actual and virtual survival functions Φ_t^{actual}, Φ_t^e, Φ_t^i, as well as weighted average $\alpha\Phi_t^e + (1-\alpha)\,\Phi_t^i$, with corresponding adjustment hazards λ_t for the GL and NS models, calibrated as in table 1.

Our expressions for Φ^e, Φ^i, and α have a sufficient statistic interpretation that may be of independent interest. In principle, given a law of motion of price gaps observed empirically, one can compute $E^t(x)$, recover Φ_t^e, Φ_t^i and α from our formulas, and therefore form the pass-through matrix from this information alone—without ever needing to solve the full model. We follow an approach of this kind in section 6.

3.3 Properties of the equivalent time-dependent models

At the heart of the equivalence result in proposition 1 are the virtual survival functions Φ_t^e, Φ_t^i. Here, we study these functions in our calibrated examples and discuss their general properties.

Figure 6 plots Φ_t^e, Φ_t^i for the GL and NS models, as well as their associated hazards. Two facts stand out. First, within each model, the (virtual) extensive and intensive hazards converge to a common limit. Second, these hazards are noticeably greater in the GL model than in the NS model. This shows that the GL model is equivalent to a mixture of TD models with shorter-lived

prices, reflecting its lower degree of monetary non-neutrality.[27]

It is tempting to compare the virtual survival and hazard functions to the actual functions that we would obtain by counting how long prices survive in panel data simulated from the SD model. Figure 6 plots these actual survival functions and hazards, constructed as the probability a price that adjusts at date 0 survives until date t without adjusting at any date $s \leq t$,

$$\Phi_t^{\text{actual}} \equiv \mathbb{P}\left(\text{no adj. until } t | x_0 = 0\right)$$

The actual SD hazards in both the GL and NS are increasing (see e.g. Alvarez and Lippi 2014) and at all times significantly below the hazards of the equivalent TD models. This implies that in both models, the aggregate price level is much more flexible than one would infer from using Φ_t^{actual} in a time-dependent model. We can prove this formally for the asymptotic hazards.

Proposition 2. *In the canonical menu cost model, the adjustment hazards λ_t^e, λ_t^i corresponding to Φ_t^e and Φ_t^i converge to the same limit $\lambda_\infty^{virtual}$. This limit is strictly above the limit of the actual adjustment hazard λ_∞^{actual}.*

This proposition extends an earlier result by Alvarez and Lippi (2022). Alvarez and Lippi (2022) showed that the asymptotic speed of convergence of the aggregate price level in response to a permanent nominal marginal cost shock is strictly above the asymptotic adjustment hazard of individual prices. Proposition 2 implies that their result holds in response to *any* shock, and that it holds separately for the responses of the extensive and intensive margins.[28]

Why do λ_t^e and λ_t^i converge to the same limit? Recall that the virtual survival functions $\Phi_t^e = E^t(\bar{x})/\bar{x}$ and $\Phi_t^i = E^{t\prime}(0) = \lim_{x\to 0} E^t(x)/x$ represent the persistence of price gaps over time, starting from the adjustment bands and reset point, respectively. Asymptotically, if there has not been a reset, ergodicity kicks in, and the rate of persistence becomes independent of the starting point.

Following Alvarez and Lippi (2022), we can attribute the gap between the asymptotic virtual and actual hazards to a selection effect. Our analytical expressions for the extensive and intensive margins shed light on selection effects for both margins. Recall that the extensive margin virtual survival curve is given by $\Phi_t^e = E^t\left(\bar{x}\right)/\bar{x}$, and that the average price gap after adjustment is zero. Hence, one way to understand why Φ_t^e differs so much from the actual survival curve Φ_t^{actual} is through the following decomposition:

$$\Phi_t^e = \Phi_t^{\text{actual}} \times \frac{\mathbb{P}\left(\text{no adj. until } t | x_0 = \bar{x}\right)}{\mathbb{P}\left(\text{no adj. until } t | x_0 = 0\right)} \times \frac{\mathbb{E}\left[x_t | \text{no adj. until } t, x_0 = \bar{x}\right]}{\bar{x}} \tag{34}$$

[27] As shown in figure 4, the GL model also has a greater weight $\alpha = 0.62$ on the extensive margin than the NS model, where $\alpha = 0.36$.

[28] Like Alvarez and Lippi (2022), we establish this by characterizing the eigenvalues and eigenfunctions of the transition operator without reinjections. In our discrete-time setting with a more general assumption on the distribution of shocks, we no longer have analytical formulas for these, but we can still show $\lambda_\infty^{actual} < \lambda_\infty^{virtual}$ by relating them to the leading eigenvalues for even and odd eigenfunctions, respectively.

The two factors on the right of (34) give the two reasons why Φ_t^e declines faster than Φ_t^{actual}. First, prices at the boundary $\bar{x}$ are less likely to survive than prices at the reset point, so the middle term is strictly below 1. Second, price gaps that do not adjust for t periods are selected: on average, they have received idiosyncratic shocks that took them closer to the middle of the Ss band, rather than pushing them outside. Thus, the term on the right in (34) is also below 1.

This discussion highlights the importance of "selection effects" in the extensive margin of price adjustment, which are well understood in the literature (e.g. Golosov and Lucas 2007). However, a similar decomposition to (34) shows that there is also a selection effect in the intensive margin of price adjustment. Decomposing $\Phi_t^i \equiv E^{t\prime}(0)$, we have:

$$\Phi_t^i = \Phi_t^{\text{actual}} \times \frac{\partial}{\partial x} \mathbb{E}\left[x_t | \text{no adj. until } t, x_0 = x\right]\Big|_{x=0} \tag{35}$$

The right factor in (35) measures the extent of this intensive margin selection effect. The logic behind this effect is the same as for the extensive margin: price gaps that do not adjust for t periods are selected, and tend to be closer to zero, the middle of the Ss band. This selection of price gaps toward zero means that following a reset, the expected persistence of the price gap (Φ_t^i) is less than the expected survival of the price itself (Φ_t^{actual}).

Indeed, as proposition 2 shows, asymptotically the extensive and intensive selection effects are equally powerful, leading to the same hazard $\lambda_\infty^{\text{virtual}} > \lambda_\infty^{\text{actual}}$. Initially, however, extensive margin selection is almost always stronger, because the probability of adjusting starting from $\bar{x}$ is much higher than starting from 0. This not only makes the middle factor of (34) below 1, but also makes the selection effect in the rightmost factor of (34) initially much stronger than in (35). In practice, this leads to the following general pattern.

Remark 1. The hazards corresponding to Φ_t^e generally increase over time; the hazards corresponding to Φ_t^i generally fall over time.

This property holds for all parameterizations of the canonical menu cost model with normal shocks we study in figure 9, as well as for the leptokurtic shocks discussed in appendix D.5. However, contrary to proposition 2, here we do not have an analytical result, and indeed we have found that the property fails for certain pathological distributions of idiosyncratic shocks.[29]

3.4 Relation to Gertler and Leahy (2008)

Gertler and Leahy (2008) is an important antecedent to our proposition 1. Gertler and Leahy (2008) gave an example of a menu cost model with a particular distribution of idiosyncratic shocks that is first-order equivalent to a Calvo model. Here, we re-derive their result in our context by showing that the two models have the same pass-through matrix, and we discuss the behavior of the extensive and intensive margin hazards in this case.[30]

[29] A counterexample can be obtained by setting $\bar{x} = 1$ and having idiosyncratic shocks distributed according to $f(x) \propto e^{-(x/1.5)^8}$. The extensive and intensive margin hazards of this model are then both non-monotonic.

[30] A previous example of a tractable menu cost model is Danziger (1999). That paper does not fit into the framework

The Gertler and Leahy (2008) example is as follows. In the canonical menu cost model, set the probability of a free adjustment to zero, $\lambda = 0$, and assume the following distribution: idiosyncratic shocks are zero with probability $1-\eta$, and are otherwise uniformly distributed in an interval $[-M, M]$. Assume further that $\eta \in (0, 1]$ and that $M > 2\overline{x}$.[31]

In this model, the expected price gap function $E^t(x)$ has a very simple shape. To see why, consider $E^1(x)$. Any price gap x in the Ss interval remains at x with probability $1-\eta$. With probability η, the idiosyncratic shock is drawn from the uniform interval, sending x to $[x-M, x+M]$. By assumption, this interval includes $[\underline{x}, \overline{x}]$. So either the price gap lands outside the Ss band, in which case it adjusts to $x^* = 0$, or it remains inside the Ss band, in which case it is uniformly distributed within $[\underline{x}, \overline{x}]$, with expectation 0. Thus,

$$E^1(x) = \underbrace{(1-\eta)\, x}_{\text{zero shock}} + \underbrace{\eta \frac{\overline{x}}{M} \cdot 0}_{\text{uniform shock, no adj.}} + \underbrace{\eta \left(1 - \frac{\overline{x}}{M}\right) \cdot 0}_{\text{uniform shock, adj.}} = (1-\eta)\, x$$

Pursuing the same logic for $t \geq 1$ shows that $E^t(x) = (1-\eta)^t x$: expected price gaps exponentially converge to zero at rate η. Using the formulas for Φ_t^e and Φ_t^i in (20), we therefore obtain:

$$\Phi_t^i = E^{t\prime}(0) = \Phi_t^e = \frac{E^t(\overline{x})}{\overline{x}} = (1-\eta)^t \tag{36}$$

Hence, the virtual survival functions are identical, and with the same constant adjustment hazard η. Applying proposition 1, we find that the pass-through matrix of the Gertler and Leahy (2008) model is identical to that of a Calvo model, with Calvo frequency η. As Gertler and Leahy (2008) point out, this virtual Calvo frequency η is higher than the actual frequency of price adjustment, $\text{freq} = \eta\left(1 - \frac{\overline{x}}{M}\right)$, delivering less monetary non-neutrality in the menu cost model than would be inferred from the frequency of price adjustment alone.

The reader may wonder if there are other examples than Gertler and Leahy (2008) in which an SD model is exactly equivalent to Calvo or to a single TD model. It turns out that the answer is yes. We give such an example in appendix C.5.

4 Numerical Equivalence between SD and Calvo Pricing Models

The Gertler and Leahy (2008) model is an important but special example in which the extensive and intensive margin hazards are exactly constant and the menu cost model is exactly equivalent to a Calvo model. This is not true more generally: instead, in typical calibrations of the canonical

of this section as it relies on a non-quadratic objective. With a quadratic approximation to the objective, it turns into the Gertler and Leahy (2008) model with $\eta = 1$, and is therefore perfectly flexible.

[31] Note that this distribution of idiosyncratic shocks does not satisfy the regularity conditions in section 2.1. It has density $f(x) = (1-\eta)\delta(x) + \frac{\eta}{2M}\mathbf{1}_{x\in[-M,M]}$, which has a mass point at $x = 0$, and also does not satisfy strict single-peakedness, or differentiability at $-M$ and M. The proof of proposition 3 itself goes through, however, since it does not require conditions on f other than symmetry and continuity at the bands $-\bar{x}$ and $\bar{x}$, which are still satisfied here.

menu cost model, extensive margin hazards are declining, and intensive margin hazards are increasing towards their common asymptotic value. Figure 6 illustrates this fact in the case of our benchmark GL and NS calibrations.

The figure also shows, however, that the hazard rate implied by the *average* virtual survival function $\alpha\Phi_t^e + (1-\alpha)\,\Phi_t^i$, plotted in the dotted black line, is, in fact, approximately constant in these examples. This suggests that these models may still effectively be close to a Calvo model, even if each of the two TD components separately is *not* close to a Calvo model.

In this section, we show that this is true across a wide range of parameterizations of the canonical menu cost model: the pass-through and Phillips curve matrices are numerically very close to those of a Calvo model. Moreover, this numerical equivalence result extends to broader menu cost models beyond the canonical model. This result is surprising, because for arbitrarily chosen TD models, both the pass-through matrix and the generalized Phillips curve are noticeably different from Calvo. This is the case, for instance, for the Taylor model discussed in appendix B.1.

4.1 Distance between pricing models

We start by defining a notion of distance between pass-through or generalized Phillips curve matrices, which will allow us to make quantitative statements about the numerical proximity of two models. For two Jacobian matrices $\mathbf{J}, \mathbf{J}'$, we define their relative distance as:

$$\text{dist}\left(\mathbf{J}, \mathbf{J}'\right) = \frac{\|\mathbf{J} - \mathbf{J}'\|}{\|\mathbf{J}\|} \tag{37}$$

where $\|\cdot\|$ is the operator norm induced by the standard L_2 norm in $\mathbb{R}^{\mathbb{N}}$.

To see why this notion of distance is natural and useful, consider first the comparison between the generalized Phillips curves of two Calvo models with slope parameters κ, κ', that is, $\mathbf{J} = \mathbf{K}^{\text{Calvo}}(\kappa)$, $\mathbf{J}' = \mathbf{K}^{\text{Calvo}}(\kappa')$. The denominator in (37) captures the average slope of the Phillips curve: how much of an inflation response a unit-standard-deviation real marginal cost shock can generate. For a Calvo model, we have[32]

$$\|\mathbf{K}^{\text{Calvo}}(\kappa)\| = \sup_{\widehat{\mathbf{mc}}:\, \|\widehat{\mathbf{mc}}\|=1} \|\mathbf{K}^{\text{Calvo}}(\kappa)\cdot\widehat{\mathbf{mc}}\| = \frac{\kappa}{1-\beta} \tag{38}$$

The numerator in (37) captures the worst-case standard deviation of the differential inflation response across the two models,

$$\|\mathbf{K}^{\text{Calvo}}(\kappa) - \mathbf{K}^{\text{Calvo}}(\kappa')\| = \sup_{\widehat{\mathbf{mc}}:\, \|\widehat{\mathbf{mc}}\|=1} \|\mathbf{K}^{\text{Calvo}}(\kappa)\cdot\widehat{\mathbf{mc}} - \mathbf{K}^{\text{Calvo}}(\kappa')\cdot\widehat{\mathbf{mc}}\|$$

[32] To get this result, consider an AR(1) shock to real marginal cost with persistence $\rho \in (0,1)$. Then, we have that $(1-\beta\rho)\,\text{sd}(\pi) = \kappa\text{sd}(\widehat{mc})$. Thus, $\|\boldsymbol{\pi}\|/\|\widehat{\mathbf{mc}}\| = \frac{\kappa}{1-\beta\rho}$ which is maximized for $\rho \to 1$, giving $\frac{\kappa}{1-\beta}$. To see that this is indeed the supremum, notice that by applying the triangle inequality to the NKPC, we have $\|\boldsymbol{\pi}\| \leq \kappa\|\widehat{\mathbf{mc}}\| + \beta\|\boldsymbol{\pi}\|$ and thus $\|\boldsymbol{\pi}\|/\|\widehat{\mathbf{mc}}\| \leq \frac{\kappa}{1-\beta}$.

One can evaluate this norm similarly to (38), finding $\|\mathbf{K}^{\text{Calvo}}(\kappa) - \mathbf{K}^{\text{Calvo}}(\kappa')\| = \frac{|\kappa-\kappa'|}{1-\beta}$. This then gives us the distance of the two Calvo models

$$\text{dist}\left(\mathbf{K}^{\text{Calvo}}(\kappa), \mathbf{K}^{\text{Calvo}}(\kappa')\right) = \frac{|\kappa - \kappa'|}{\kappa}$$

Intuitively, our measure of distance in (37) captures the relative difference in Phillips curve slopes. In the following, we apply (37) to compute the distance between the generalized Phillips curve $\mathbf{J} = \mathbf{K}$ of a menu cost model and the generalized Phillips curve of a Calvo model, $\mathbf{J}' = \mathbf{K}^{\text{Calvo}}(\kappa)$.

In principle, the distance measure (37) can also be used to compare pass-through matrices. Since pass-through matrices and generalized Phillips curves are related by the one-to-one mapping in (16), the distance measures end up being very similar. However, the distance between Phillips curve matrices has a more intuitive natural interpretation in terms of relative slopes, so we take it as our benchmark measure. We consider alternative distance measures in robustness checks.

4.2 Numerical equivalence result

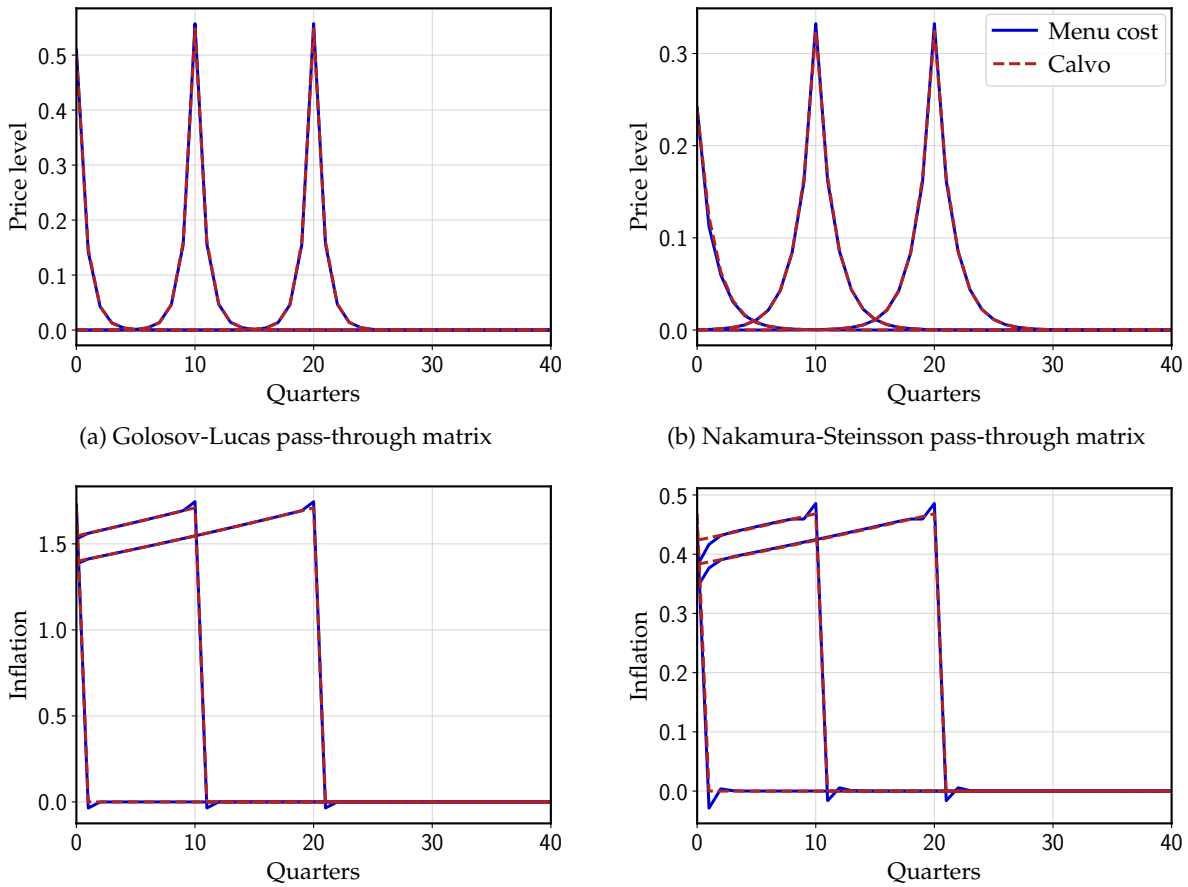
The blue lines in figure 7 show the columns of the pass-through and Phillips curve matrices of the GL and NS models. As expected, the GL pass-through matrix is more spiked around the date of the anticipated nominal marginal cost shock than the NS pass-through matrix, indicating that the GL model is closer to flexible prices (see our discussion in section 2.3). Accordingly, the GL generalized Phillips curve has much larger columns than the NS generalized Phillips curve: the same-sized real marginal cost shock increases inflation by more than three times as much in GL relative to NS.

The red dashed lines in figure 7 show the best Calvo approximations obtained by minimizing the distance measure (37) for each of the two models. The fit is very close. The only visible deviations arise in early periods for the NS model, and in both models in periods around the date of the real marginal cost shock. The best fitting Calvo models for GL and NS have hazards λ of 0.707 and 0.487. Given $\beta = 0.99$, this implies Calvo slopes κ of 1.709 and 0.468, respectively.

Remark 2. The GL and NS models are numerically first-order equivalent to Calvo.

The fact that a Calvo model can fit the aggregate pricing behavior of SD models so well is surprising. For instance, it is well known that menu cost models have upward sloping adjustment hazards (see for instance Alvarez and Lippi 2014, and the green line in figure 6). It is also well known that, in a time dependent model, upward sloping adjustment hazards imply an inertial Phillips curve (see Sheedy, 2010, and figures 1-3, panel b). Combining these results, it would be natural to expect SD models to also feature inflation inertia. Yet they do not: a real marginal cost shock in the SD models in figure 7 neither causes a slow build-up in inflation initially, nor causes a slow reduction in inflation after the time of the shock.

The reason for this lack of inertia is the difference between virtual and actual hazards. As figure 6 shows, in both the GL and the NS models, actual hazards increase, but average virtual



(a) Golosov-Lucas pass-through matrix

(b) Nakamura-Steinsson pass-through matrix

(c) Golosov-Lucas generalized Phillips curve

(d) Nakamura-Steinsson generalized Phillips curve

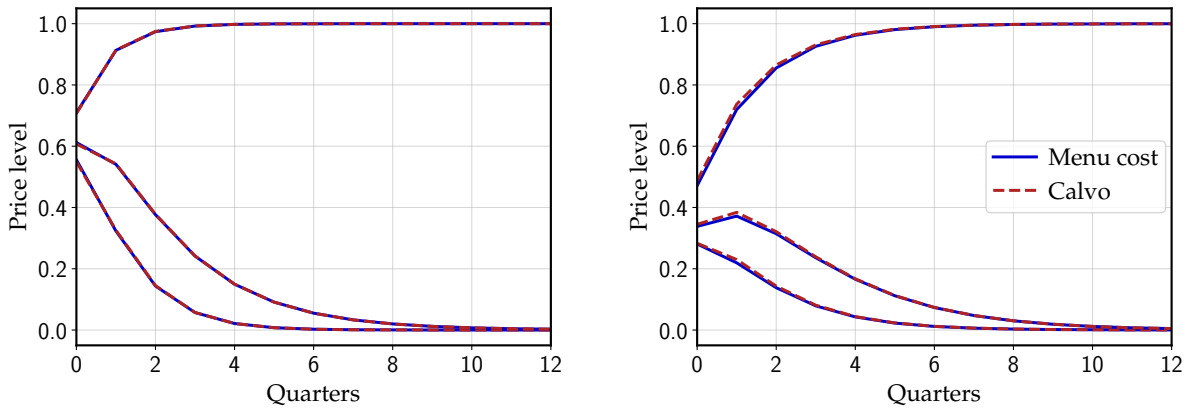
Figure 7: Approximating menu cost Phillips curves with Calvo models

Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the GL and NS models, calibrated as in table 1, as well as the best-fitting Calvo approximations, obtained using the distance measure (37) applied to Phillips curve matrices.

hazards do not. In other words, the selection effect undoes the inertia in the Phillips curve that one would expect from actual hazards.

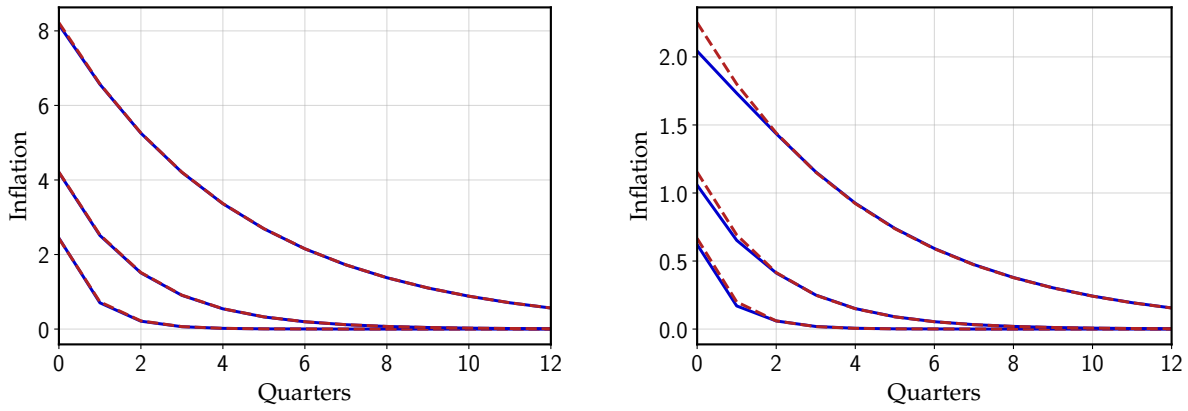
Relationship to earlier equivalence results. Several papers have previously explored the difference between SD and Calvo models. Our numerical equivalence result significantly extends these earlier findings. Alvarez et al. (2016) and Alvarez et al. (2017) characterize the cumulative impulse response (CIR) of one minus the price level (which is output in their model) to a unit-sized permanent shock to nominal costs in menu cost vs Calvo models, expressing those in terms of the kurtosis of the stationary distribution of price changes. Alvarez and Lippi (2022) characterize the entire impulse response of the price level to permanent shocks to nominal costs by finding the eigenvalues and eigenfunctions of the relevant dynamical system. One of their findings is that this particular impulse response is well approximated by that of a Calvo model.

By contrast, our numerical equivalence result establishes that the entire impulse response to *ar-*



(a) Golosov-Lucas nominal shocks.

(b) Nakamura-Steinsson nominal shocks.



(c) Golosov-Lucas real shocks

(d) Nakamura-Steinsson real shocks

Figure 8: Price level responses to AR(1) marginal cost shocks

Note: impulse responses to AR(1) marginal cost shocks for GL and NS models, calibrated as in table 1, as well as the best-fitting Calvo approximations, obtained using the distance measure (37). Shock persistence values are $\{0.3, 0.6, 1\}$ for nominal shocks (top) and $\{0.3, 0.6, 0.8\}$ for real shocks (bottom).

bitrary nominal or *real* marginal cost shocks is well approximated by a Calvo model. This extension to all shocks is important, since it shows that the two price-setting models are effectively the same without restrictive assumptions on preferences or on the nature of aggregate shocks. We illustrate this result in figure 8 for various processes for nominal and real marginal cost shocks. The close match across responses follows directly from the fact that both the pass-through and Phillips curve matrices are sufficient statistics for the aggregate pricing behavior of a state-dependent model. If a Calvo matches these matrices well, it matches the entire first-order aggregate behavior of the SD model, including impulse responses to all shocks.

An interesting observation from figure 8 is that the Calvo approximation works somewhat less well for the NS model than for the GL model, in spite of the prevalence of free adjustments. The reason is as follows. While both models are calibrated to have the same frequency of price changes, in the GL model these adjustments are entirely triggered by price gaps leaving the inaction region, thus requiring a relatively tight *Ss* band. This leads to faster mixing of price gaps, and hence faster

convergence of the intensive and extensive margin hazards, as is clear from figure 6. In turn, this makes the GL model more "Calvo-like".

By contrast, the NS model has a significant share of free adjustments. Holding the overall frequency of adjustment constant, this requires a wider *Ss* band, to reduce the frequency of extensive margin adjustments. But the wider *Ss* band reduces the speed of mixing between price gaps, and the speed of convergence of the intensive and extensive margin hazards in figure 6. This, paradoxically, makes the NS model less "Calvo-like". We illustrate this result in figure C.3 of appendix C.7, where we show that the extensive margin hazard converges very slowly for shares of free adjustments close to 1.

Estimating the NKPC on data from the menu cost model. The numerical equivalence between SD and Calvo models has a simple implication: menu cost models are well described by the NKPC, for a model-specific slope parameter κ. This suggests an alternative distance metric: one can simulate data from an SD model, estimate (NK-PC) on the simulated data, and use the R^2 from the regression as a measure of fit. Implementing this procedure in the GL and NS models using an AR(1) process for real marginal cost with quarterly persistence 0.8 delivers $R^2 = 1.000$ (GL) and $R^2 = 0.998$ (NS), also suggesting a tight match overall. Going beyond this procedure, we also estimate a hybrid Phillips curve, as in Galí and Gertler (1999). This delivers an estimated term on lagged inflation that is very close to 0, confirming the fact that menu cost models cannot generate inflation inertia. Details are provided in appendix D.1.

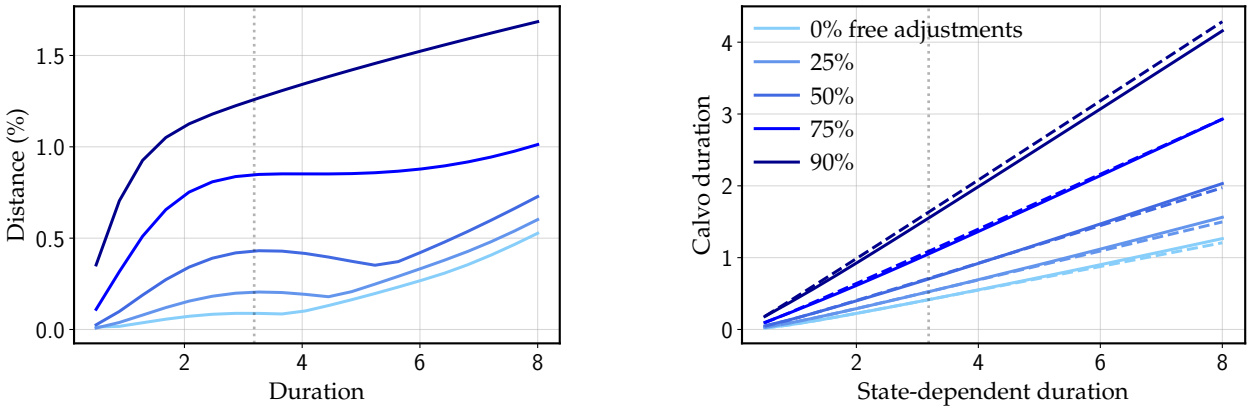
4.3 Robustness to parameterization of the canonical menu cost model

The GL and NS models are only two parametrizations of the canonical menu cost model. Here, we systematically explore the two-dimensional parameter space of this model within the class of normal idiosyncratic shocks.[33]

Figure 9(a) plots the relative distance (37) between the generalized Phillips curve and its best Calvo fit, as we vary both the duration (defined as $\frac{1}{\text{freq}} - 1$) and the share of free price adjustments. The dotted vertical line indicates the empirical duration in our baseline GL and NS calibration; recall that our GL calibration has no free adjustments and our NS calibration has a share of free adjustments of 75%.

At smaller durations, we find a smaller distance, indicating a closer match between the SD and Calvo models. This is intuitive as tighter *Ss* bands not only increase frequency, but also increase the speed of mixing of price gaps, leading the intensive and extensive margin hazards to converge more quickly. This brings the model closer to Calvo. Less intuitively, but in line with our discussion of the worse Calvo fit of the NS model relative to GL, a greater share of free adjustments can increase the distance to Calvo. The distance only falls to zero for adjustment shares very close to 100% (not shown).

[33] In appendix D.5 we consider an extension with leptokurtic shocks.



(a) Distance

(b) Distance-minimizing (solid) Calvo duration and the one implied by Alvarez et al. (2016) (dashed)

Figure 9: Varying the duration and share of free adjustments

Note: the left panel shows the distance between the state-dependent Phillips curve and the best-fitting Calvo approximation, measured with distance (37). The right panel shows the duration of the Calvo minimizer, as well as the one implied by the Alvarez et al. (2016) result, that is, $d^{Calvo} = \frac{1}{6}\text{Kur}^{\text{MC}}\left(1 + d^{MC}\right) - \frac{1}{2}$ for the value of Kur^{MC} in the simulation of the menu cost model under the duration d^{MC} and the share of free adjustment under consideration. The vertical dotted lines show the empirical duration that we target in our baseline calibration.

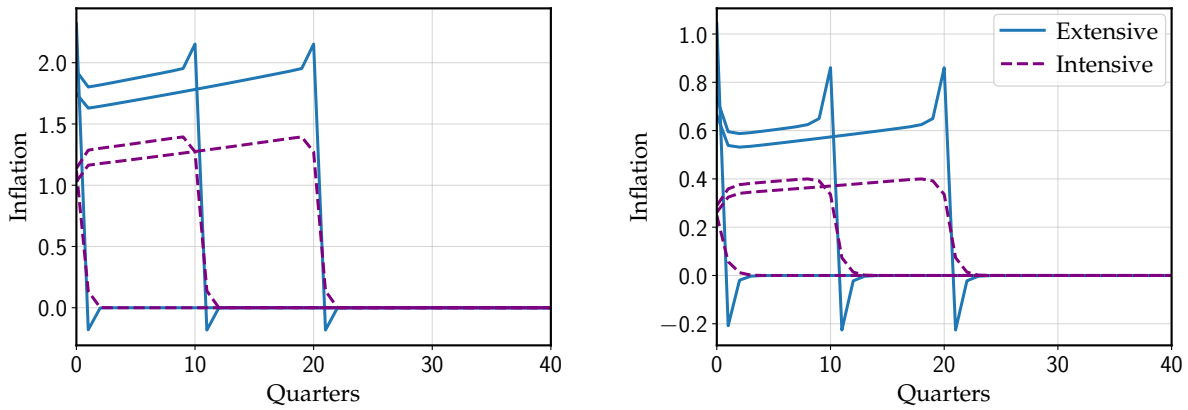
4.4 Why do Calvo and menu cost models have such close aggregate predictions?

In the beginning of section 4, we observed that the average virtual survival function $\alpha\Phi_t^e + (1-\alpha)\,\Phi_t^i$ was close to a Calvo survival function $\Phi_t^{\text{Calvo}} = (1-\lambda)^t$, and conjectured that as a result, aggregate behavior might be close to Calvo as well. We then verified this conjecture numerically. In this section, we elaborate on the reasons for this close numerical fit.

In particular, we show in appendix D.2 that if we write the virtual survival functions as $\Phi_t^e = \Phi_t^{\text{Calvo}} + \eta_t^e$ and $\Phi_t^i = \Phi_t^{\text{Calvo}} + \eta_t^i$, then to a first-order approximation in the ηs, the gap between the actual pass-through matrix and the Calvo pass-through matrix scales with $\alpha\Phi_t^e + (1-\alpha)\Phi_t^i - \Phi_t^{\text{Calvo}}$. In other words, if the virtual survival functions are not individually too far from Calvo, and their average is close to Calvo, then the pass-through matrix—and consequently the generalized Phillips curve, via (16)—will be close to Calvo as well.

In figure 6, particularly for the GL model, these conditions are satisfied, explaining why Calvo is such a good approximation: the extensive and intensive survivals are not far from Calvo, and their mixture is even closer. By proposition 2 and remark 1, we expect this to be true quite generally: the two hazard rates always converge to the same constant, and prior to convergence they deviate in offsetting directions.

It is important that the deviations from Calvo in the extensive and intensive survival functions offset each other when averaged, since individually these deviations are larger. In figure 10, we separately plot the generalized Phillips curves corresponding to the intensive and extensive margin time-dependent models. We see that the two margins separately produce fairly different generalized Phillips curves, deviating from the New Keynesian Phillips curve by far more than



(a) Golosov-Lucas. (b) Nakamura-Steinsson.

Figure 10: Generalized Phillips curves associated with intensive and extensive margins

Note: Generalized Phillips curves for intensive and extensive pass-through matrices in GL and NS models, calibrated as in table 1.

in figure 6. In particular, there is anti-persistence in the extensive margin and persistence in the intensive margin. This difference is most pronounced for the NS model.

4.5 Forward-lookingness in the generalized Phillips curve

Although the generalized Phillips curves in figure 10 are distinct from the NKPC, one similarity is striking: as we go backwards in time from the quarter of the shock, both of these margins appear to be decaying at the same rate β as the Calvo Phillips curve. As the following proposition shows, this turns out to be a general result for both time-dependent and menu cost models.[34]

Proposition 3. *Let* $\mathbf{K} = (K_{t,s})$ *be the generalized Phillips curve of an arbitrary convex combination of TD or canonical menu cost pass-through matrices. Then, the columns of* $\mathbf{K}$ *converge to a two-sided sequence* $\{k_j\}$ *around the diagonal, i.e.* $K_{s+j,s} \to k_j$ *as* $s \to \infty$*, for each* $j \in \mathbb{Z}$*. Going backward, this sequence decays at rate* β *asymptotically, i.e.*

$$\lim_{j\to\infty} \frac{k_{-j}}{\beta^j} = C$$

for some constant C*.*

This proposition shows that the extreme "forward-lookingness" of the NKPC, with future shocks discounted at rate β irrespective of the horizon—which plays an major role in the forward guidance puzzle (Del Negro et al., 2023)—is present in all price-setting models we have introduced in this paper.

[34]For this proposition, we will impose the regularity condition that survival rates of the TD models in question eventually decline exponentially, i.e. that $\Phi_s < C/v^s$ for some constant C and $v > 1$.

4.6 Which Calvo frequency fits best?

So far, we have recovered the Calvo model that most closely approximates a given menu cost model by simulating the latter. Here, we show that it is possible to avoid simulation, and instead use a result by Alvarez et al. (2016) to directly recover the slope κ.

Figure 9(b) shows the Calvo duration that provides the best fit to each model across the parameter space of the canonical menu cost model. From any Calvo duration d, we can back out the equivalent Calvo frequency $\text{freq} = 1/(1+d)$, and therefore the slope κ using the standard formula,

$$\kappa = \frac{\text{freq}\,(1-\beta\,(1-\text{freq}))}{1-\text{freq}} = \frac{1-\beta\frac{d}{1+d}}{d} \tag{39}$$

Observe that the relation between menu cost duration and Calvo duration in figure 9(b) is close to linear. This is an instance of the Alvarez et al. (2016) result. Alvarez et al. (2016) showed that, in continuous time, the cumulative impulse response of the price level to a permanent nominal cost shock depends only on the ratio of the kurtosis of price changes to the frequency. Because the Calvo model provides a close approximation to the menu cost model, it must approximately have the same CIR, and therefore the same ratio of kurtosis to frequency,

$$\frac{\text{Kur}^{\text{Calvo}}}{\text{freq}^{\text{Calvo}}} \simeq \frac{\text{Kur}^{\text{MC}}}{\text{freq}^{\text{MC}}} \tag{40}$$

In continuous time, $\text{Kur}^{\text{Calvo}}$ is equal to 6, and Kur^{MC} is a constant less than 6 that depends only on the share of free adjustments, and is equal to 1 in the Golosov-Lucas case. Hence, (40) implies a linear relationship between Calvo and menu cost duration, with a slope equal to $\frac{1}{6}$ under zero free adjustments and increasing as the share of free adjustment rises. These properties are all apparent in figure 9(b). The dashed lines provide a numerical approximation based on the simulated values of $\text{Kur}^{\text{Calvo}}$ and Kur^{MC} in the discrete-time model (where these values are no longer independent of frequency) and show that this formula provides an excellent quantitative fit.[35]

In conclusion, equations (39) and (40), combined with the relationship between Calvo kurtosis and frequency, allow us to obtain a Phillips curve based entirely on the ratio of kurtosis to frequency in the menu cost model.[36] In the limit as $\beta \to 1$, this becomes:

$$\kappa \simeq \frac{4}{\left(\frac{1}{3}\frac{\text{Kur}^{MC}}{\text{freq}^{MC}}\right)^2 - 1} \tag{41}$$

The quantitative implications of (41) are worth stressing. In the data, the frequency of price changes is usually found to be around 0.2 and 0.3 at the quarterly frequency, and the kurtosis of price changes is typically thought to be between 3 and 4. Implementing (41) with these val-

[35] In the discrete-time Calvo model, we have $\text{Kur}^{\text{Calvo}} = 3\left(2 - \text{freq}^{\text{Calvo}}\right)$, while in the discrete-time menu cost model there is no analytical expression and Kur^{MC} must be simulated numerically.

[36] This corresponds to the value of κ when $\frac{1}{\text{freq}^{\text{Calvo}}} = \frac{1}{6}\frac{\text{Kur}^{\text{MC}}}{\text{freq}^{\text{MC}}} + \frac{1}{2}$.

ues delivers $\kappa \in [0.1; 0.4]$. By contrast, our menu cost models generate kurtosis of 1.3 for GL and 2.3 for NS at our calibrated frequency of 0.239. Applying (41), we obtain a κ of 1.75 for GL and 0.53 for NS, close to the values delivered by the best-fitting Calvo model. These values are high compared to the $\kappa = 0.08$ implied by a Calvo model with this frequency of price change, and also high compared to typical direct estimates of κ from macro data. We return to this point in the conclusion.

4.7 Extensions, trend inflation, and large shocks

In appendix D.5, we show that the numerical equivalence result between the canonical model and Calvo appears to be robust by considering several extensions to our analysis.

First, we consider first-order shocks in alternative menu cost models. In particular, we consider models with leptokurtic shocks (for which we can show that proposition 1 still applies), as well as two-product models as in Midrigan (2011). The first-order aggregate behavior of these models remains to close that of a suitably parametrized Calvo model. We also consider the multi-sector menu cost model of Nakamura and Steinsson (2010), and show that its aggregate behavior is closely approximated by a multi-sector Calvo model.

Second, we extend the GL and NS models to allow for trend inflation. We find that for modest levels of trend inflation (up to 5%), the Phillips curve of a Calvo model *without* trend inflation still provides a close approximation. In the GL and NS models, trend inflation increases the slope of the Phillips curve (see figure D.3 in appendix D.5). This is in line with existing results in Gagnon (2009), Alvarez and Neumeyer (2020) and Alexandrov (2022). By contrast, as we show in appendix E.7, in a Calvo model, trend inflation reduces the slope of the Phillips curve, and also changes its shape. As a result, the Calvo model with trend inflation is actually a poor approximation to the menu cost model with trend inflation, and the traditional Calvo NKPC performs better.[37]

Finally, we consider large shocks in the canonical menu cost model. In appendix D.5, we show that the responses to shocks of up to 5% to nominal marginal cost (figure D.7), or shocks to real marginal cost that cause inflation of up to 5% (figure D.8), are still well approximated by the Calvo model. By contrast, shocks larger than 10% lead to greater departures from numerical equivalence with reduced monetary non-neutrality, in line with previous literature on menu cost models (Karadi and Reiff 2019, Bonadio, Fischer and Sauré 2020, Auer, Burstein and Lein 2021, Alexandrov 2022, Blanco, Boar, Jones and Midrigan 2022, Hall 2023).

Figure D.8 also explores aggregate state dependence, that is, whether large past shocks can influence the impulse responses to additional shocks later on. We find limited evidence of such effects in the context of our calibrations of the canonical menu cost model, provided that shocks remain below 5%.

[37] One intuition is that key forces in the Calvo model with trend inflation—for instance, the fact that a few firms will have very low prices and high demand because they have not adjusted for a long time—disappear in menu cost models, where there is always the option to adjust. Our results on trend inflation relate to the literature on "state dependence" (as in Vavra 2014), where the level of underlying trend inflation determines the overall degree of monetary non-neutrality. Trend inflation can also cause sign dependence (Tsiddon 1993, Ball and Mankiw 1994).

5 General Equilibrium

So far, we have set up both state- and time-dependent models assuming a quadratic objective and linear aggregation. In the analytical menu cost literature, this is sometimes taken as a primitive environment for convenience, and is usually viewed as the correct approximation to a deeper microfounded price-setting problem (e.g. Alvarez and Lippi 2014). We have also studied the generalized Phillips curve, solving for inflation as a function of real marginal cost in (16), by analogy to the New Keynesian Phillips curve.

In this section, we justify the use of both the quadratic approximation and the generalized Phillips curve in the context of fully microfounded general equilibrium DSGE models with menu cost price-setting.[38] We show that the first-order perturbation solution of this model is, as idiosyncratic risk becomes small, exactly the same as that of the same model with the generalized Phillips curve **K** replacing the entire price-setting model. We formally show this first in the context of the standard New Keynesian model, with and without strategic complementarity, and then discuss how the result extends to a more complex DSGE model.[39]

5.1 Textbook New Keynesian model with menu costs

Our model is set in discrete time. We closely follow Galí (2008) in terms of model structure and notation, except for the price-setting behavior of the firm. We continue to write the model under perfect foresight over aggregate variables, and to denote log-deviations from the steady state with a hat.

Households. The model is populated by a representative household maximizing the utility function

$$\sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - b \frac{N_t^{1-\varphi}}{1+\varphi} \right]$$

over paths of consumption and hours $\{C_t, N_t\}$ subject to the flow budget constraint

$$P_t C_t + B_t \leq (1 + i_{t-1}) B_{t-1} + W_t N_t + P_t \Pi_t$$

and a standard No-Ponzi condition $\lim_{T \to \infty} B_T \geq 0$. The first-order conditions of this problem are given by the usual expressions

$$C_t^{-\sigma} = \frac{P_t}{P_{t+1}} (1 + i_t) C_{t+1}^{-\sigma} \qquad \text{and} \qquad b N_t^{\varphi} = \frac{W_t}{P_t} C_t^{-\sigma} \tag{42}$$

[38] Earlier examples of DSGE models with menu cost pricing are Dotsey et al. (1999) and Costain and Nakov (2011).

[39] As pointed out by Fernández-Villaverde (2010), general equilibrium menu cost models are hard to simulate: "The bad news is, of course, that handling a state-dependent pricing model is rather challenging (we have to track a non-trivial distribution of prices), which limits our ability to estimate it. Being able to write, solve, and estimate DSGE models with better pricing mechanisms is, therefore, a first order of business."

Consumption C_t is an index bundling many varieties i,

$$C_t \equiv \left(\int_0^1 \left(\frac{C_{it}}{A_{it}} \right)^{\frac{\zeta-1}{\zeta}} di \right)^{\frac{\zeta}{\zeta-1}} \tag{43}$$

where $\zeta > 1$ is the elasticity of substitution between varieties; and A_{it} are idiosyncratic preference shifters. We define aggregate output to be equal to consumption, $Y_t = C_t$. Demand for variety i and the aggregate price index are then given by

$$Y_{it} = A_{it}^{1-\zeta} \left(\frac{P_{it}}{P_t} \right)^{-\zeta} Y_t \qquad \text{and} \qquad P_t = \left(\int_0^1 (A_{it} P_{it})^{1-\zeta} di \right)^{\frac{1}{1-\zeta}} \tag{44}$$

where P_{it} denotes the price of variety i.

Firms. There is a continuum of monopolistically competitive firms. Firm i produces quantity Y_{it} of variety i with linear production function $Y_{it} = A_{it} N_{it}$ from hours N_{it}. Importantly, variety i's preference shifter A_{it} is also firm i's productivity shock. $\log A_{it}$ evolves according to a random walk,

$$\log A_{it} = \log A_{it-1} + \sigma_\epsilon \epsilon_{it}$$

where ϵ_{it} has density $\bar{f}$ satisfying the same restrictions as in section 2. Firm i's real profits at date t are given by

$$\Pi_{it} = \frac{P_{it}}{P_t} Y_{it} - \frac{W_t}{P_t} N_{it} = \left(\frac{P_{it}}{P_t} - \frac{W_t}{P_t} \frac{1}{A_{it}} \right) \cdot A_{it}^{1-\zeta} \left(\frac{P_{it}}{P_t} \right)^{-\zeta} Y_t \tag{45}$$

The firm's statically optimal price, which we denote by $P_{it}^* W_t$ (separating into idiosyncratic P_t^* times the aggregate W_t), is therefore given by the usual constant markup rule

$$P_{it}^* W_t \equiv \frac{\zeta}{\zeta - 1} \frac{W_t}{A_{it}} \tag{46}$$

Substituting out A_{it} from (45) using (46), we can express profits as

$$\Pi_{it} = \left(\frac{\zeta}{\zeta - 1} \frac{W_t}{P_t} \right)^{1-\zeta} Y_t \cdot \left(\left(\frac{P_{it}}{P_{it}^* W_t} \right)^{1-\zeta} - \frac{\zeta - 1}{\zeta} \left(\frac{P_{it}}{P_{it}^* W_t} \right)^{-\zeta} \right) \tag{47}$$

As in section 2, we define firm i's idiosyncratic price gap as

$$x_{it} = \log P_{it} - \log P_{it}^*$$

With this notation, profits can be written entirely as a function of the price gap and aggregate variables,[40]

$$\begin{aligned}\Pi_{it} &= \left(\frac{\zeta}{\zeta-1}\frac{W_t}{P_t}\right)^{1-\zeta} Y_t \cdot \left(e^{(1-\zeta)(x_{it}-\log W_t)} - \frac{\zeta-1}{\zeta}e^{-\zeta(x_{it}-\log W_t)}\right) \\ &\equiv \left(\frac{\zeta}{\zeta-1}\frac{W_t}{P_t}\right)^{1-\zeta} Y_t \cdot F\left(x_{it}-\log W_t\right) \end{aligned} \tag{48}$$

where we have introduced the function $F(x) \equiv e^{(1-\zeta)x} - \frac{\zeta-1}{\zeta}e^{-\zeta x}$. F has a local maximum at 0, that is, $F'(0) = 0$ and $F''(0) < 0$. We can also express the price level P_t from (44) in terms of price gaps,

$$P_t = \frac{\zeta}{\zeta-1}\left(\int_0^1 e^{(1-\zeta)x_{it}}di\right)^{\frac{1}{1-\zeta}} \tag{49}$$

Inflation is now defined as $\pi_t = P_t/P_{t-1} - 1$.

As in section 2, we assume that firms have to pay a random menu cost $\bar{\xi}_{it} \in \{0, \bar{\xi}\}$ when changing their prices, where the probability of a free adjustment, $\bar{\xi}_{it} = 0$, is parametrized by $\lambda \in [0,1)$, and $\bar{\xi} > 0$. Following Golosov and Lucas (2007) and Nakamura and Steinsson (2010), we assume the menu cost are stated in units of labor required to change prices. Moreover, we scale the menu cost by σ_ϵ^2, so that the model is well behaved in the limit of small σ_ϵ. Given this, firm i's profit maximization problem reads

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t C_t^{-\sigma}\left[\left(\frac{\zeta}{\zeta-1}\frac{W_t}{P_t}\right)^{1-\zeta} Y_t \cdot F\left(x_{it}-\log W_t\right) + \sigma_\epsilon^2 \bar{\xi}_{it}\frac{W_t}{P_t} 1_{\{x_{it}\neq x_{it-1}-\sigma_\epsilon\epsilon_{it}\}}\right] \tag{50}$$

where $\beta^t C_t^{-\sigma}$ is the representative agent's stochastic discount factor up to a multiplicative constant.

The aggregate amount of labor required for menu costs is given by

$$\Xi_t \equiv \int_0^1 \sigma_\epsilon^2 \bar{\xi}_{it} 1_{\{x_{it}\neq x_{it-1}-\sigma_\epsilon\epsilon_{it}\}}di \tag{51}$$

Aggregate profits $\Pi_t \equiv \int_0^1 \Pi_{it}di$ are paid directly to the representative agent. With the help of (49), aggregate labor demand by firms can be written as

$$N_t^d \equiv Y_t\Delta_t + \Xi_t \tag{52}$$

where $\Delta_t \equiv \left(\int_0^1 e^{(1-\zeta)x_{it}}di\right)^{\frac{\zeta}{1-\zeta}} \int_0^1 e^{-\zeta x_{it}}di \geq 1$ captures the productivity loss due to price dispersion.

Monetary policy rule. We assume the central bank operates a standard Taylor rule, $i_t = \rho + \phi\pi_t + v_t$ where $\rho = \beta^{-1} - 1$ is the discount rate, $\phi > 1$, and v_t is a monetary policy shock.

[40] Without introducing A_{it} as simultaneous preference and technology shocks, this would not be feasible.

Market clearing and equilibrium. The goods market clearing condition is simply given by $C_t = Y_t$. Labor market clearing is given by $N_t = N_t^d$. Asset market clearing is given by $B_t = 0$. A competitive equilibrium is an allocation $\{C_t, N_t, N_t^d, Y_t, B_t, \Xi_t, Y_{it}, N_{it}, \Pi_t\}$ together with prices $\{P_t, P_{it}, W_t, \pi_t, i_t\}$ such that the representative agent maximizes utility, the central bank follows its rule, and all firms maximize the present discounted value of their profits.

Steady state with no aggregate shocks. A steady-state equilibrium with no aggregate shocks ($\nu_t \equiv 0$) is characterized by a set of constant aggregates $\{N_{ss}, Y_{ss}, \Xi_{ss}, P_{ss}, W_{ss}, i_{ss}, \Delta_{ss}\}$. It follows from the monetary policy rule and steady-state Euler equation (42) that steady-state inflation must be zero and $\beta(1 + i_{ss}) = 1$. We resolve indeterminacy of steady-state prices and wages with the normalization $W_{ss} = 1$.[41]

In the case with no idiosyncratic shocks or menu costs ($\sigma_\epsilon = 0$), all price gaps are zero, and we have $\Delta_{ss} = 1$, $\Xi_{ss} = 0$, $P_{ss} = \frac{\zeta}{\zeta-1}$, $Y_{ss} = N_{ss}$, and $bN_{ss}^\varphi = \frac{\zeta}{\zeta-1} Y_{ss}^{-\sigma}$. In appendix E, we show that all steady-state aggregates converge to these $\sigma_\epsilon = 0$ levels as we take the limit $\sigma_\epsilon \to 0$.

First-order response to aggregate shocks around steady state. Following a vast literature (see, in particular, Reiter 2009), we are interested in the first-order perturbation solution in aggregates around the steady state described above. In particular, we consider an arbitrary bounded perturbation $\{d\nu_t\}_{t=0}^\infty$ to the intercept of the Taylor rule from date 0 onward, assuming that the economy begins in the steady state, and we solve for the implied perturbations to endogenous variables, e.g. $\{dY_t\}$, $\{d\pi_t\}$, and $\{di_t\}$.

In appendix E.1, we describe the equations that characterize this solution in the sequence space. We note that this solution depends on the σ_ϵ that scales idiosyncratic risk. The following proposition, however, shows that in the limit of small idiosyncratic risk, the impulse responses of $\hat{Y}_t \equiv dY_t/Y_{ss}$, $\hat{\pi}_t \equiv d\pi_t$, $\hat{i}_t \equiv di/(1 + i_{ss})$, and $\hat{\nu}_t \equiv d\nu_t$ satisfy a simple analog to the standard three-equation New Keynesian model.[42]

Proposition 4. *As $\sigma_\epsilon \to 0$, the equations characterizing $\{\hat{Y}_t, \hat{\pi}_t, \hat{i}_t, \hat{\nu}_t\}$ converge to*

$$\hat{Y}_t = \hat{Y}_{t+1} - \frac{1}{\sigma}(\hat{i}_t - \hat{\pi}_{t+1}) \tag{53}$$

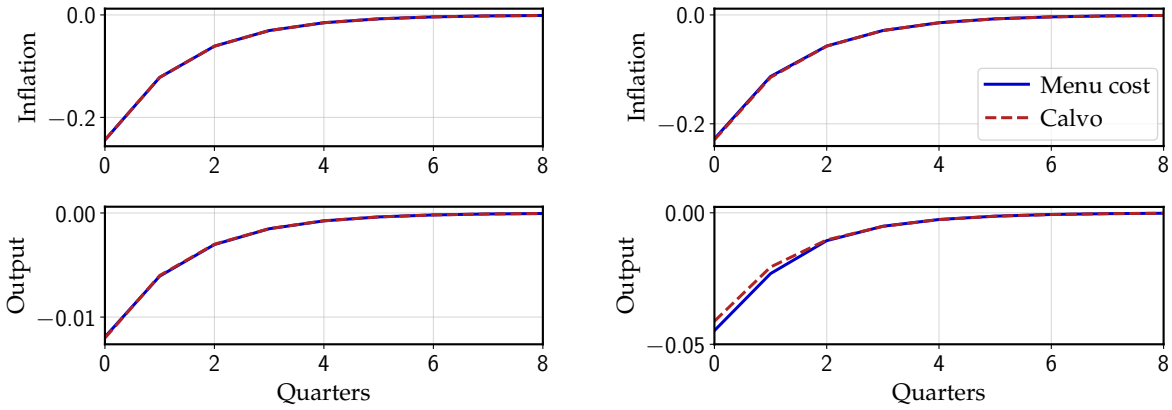
$$\hat{\boldsymbol{\pi}} = (\varphi + \sigma)\, \mathbf{K} \cdot \hat{\mathbf{Y}} \tag{54}$$

$$\hat{i}_t = \phi \hat{\pi}_t + \hat{\nu}_t \tag{55}$$

where $\mathbf{K}$ is the generalized Phillips curve implied by the canonical menu cost model in section 2.1, given the probability of free adjustments λ, idiosyncratic innovations to x/σ_ϵ distributed as $\bar{f}$, and a ratio of menu cost to idiosyncratic risk $\frac{\bar{\xi}}{\sigma_\epsilon^2} = \left(\frac{\zeta-1}{\zeta}\right)^{1-\zeta} \left(\frac{W_{ss}}{P_{ss}}\right)^{\zeta} \frac{1}{Y_{ss}} \frac{2\bar{\xi}}{F''(0)}$.

[41] The remaining equations characterizing the steady state are the labor-consumption FOC $bN_{ss}^\varphi = \frac{W_{ss}}{P_{ss}} Y_{ss}^{-\sigma}$, labor demand plus market clearing $N_{ss} = Y_{ss}\Delta_{ss} + \Xi_{ss}$, and three equations for P_{ss}, Δ_{ss}, and Ξ_{ss} from the price-setting problem.

[42] Here, the ss subscripts refer to the steady state with no idiosyncratic or aggregate risk.



(a) Golosov-Lucas (b) Nakamura-Steinsson

Figure 11: Impulse responses for the standard New Keynesian model with state-dependent pricing

Note: both pricing blocks were calibrated as in table 1. The approximating Calvo models are constructed in section (4). The other dynamics are given by equations (53) — (55).

Relative to the standard three-equation New Keynesian model, the only change in (53)–(55) is that the New Keynesian Phillips curve, which in this context is $\hat{\pi}_t = (\varphi + \sigma)\kappa\hat{Y}_t + \beta\hat{\pi}_{t+1}$, has been replaced by the generalized Phillips curve (54). As explained in section 2.5, the generalized Phillips curve **K** depends on the probability of a free adjustment λ, the distribution $\bar{f}$ of idiosyncratic shocks, and the ratio of menu cost to idiosyncratic risk ξ/σ_ϵ^2; the proposition shows how to obtain these from the primitives of the nonlinear model. In summary, proposition 4 shows that, for small enough idiosyncratic risk, the entire pricing side of the model can be summarized by (54) for the purpose of characterizing first-order aggregate impulse responses.

The intuition for proposition 4 is as follows. In most respects, our model is identical to the standard New Keynesian framework, leading to the same intertemporal Euler equation, and real marginal cost $\widehat{mc}_t = (\varphi + \sigma)\hat{Y}_t$. The key difference is that Calvo pricing is replaced by a more complex menu cost model, which leads to several complications. For instance, for $\sigma_\epsilon > 0$, both price dispersion and aggregate menu costs are time-varying and enter into the log-linearized aggregate equations; also, first-order changes in the real wage W_t/P_t and level of production Y_t enter into the firm decision problem and affect aggregate pricing. All these terms, however, vanish from the first-order aggregate system as σ_ϵ goes to zero. The only aggregate relationship that remains is between aggregate wages and aggregate prices, just as in the canonical model. Moreover, for small σ_ϵ, the firm profit objective (48) becomes quadratic and price aggregation (49) becomes linear, both as in the canonical model, leading to the same generalized Phillips curve **K**.

Figure 11 implements (53)–(55), plotting the response of the model to an AR(1) monetary policy shock v with magnitude 0.25 on impact and persistence 0.5. The calibration used is the same as that in Galí (2008): $\sigma = 1, \varphi = 5, \phi = 1.5, \rho = 0.01$. As expected, the NS model predicts an output response that is about three times as large as the one predicted by GL. Both impulse responses are closely matched when (54) is replaced by the New Keynesian Phillips curve for the best-fitting

Calvo approximation found in section 4.[43]

In appendix E.4, we show that (53)–(55) continue to provide an excellent approximation to the fully nonlinear case even when σ_ϵ is set to match our original calibration, which targets the average size of price changes.

5.2 Strategic complementarities

Both output responses in figure 11 are relatively modest. As pointed out in the literature, this partly reflects the lack of strategic complementarities.

Following Caplin and Leahy (1997) and Nakamura and Steinsson (2010), we now introduce strategic complementarities to the model by assuming roundabout production of a particular type.[44] We modify firm i's production function to be $Y_{it} = A_{it} N_{it}^{\chi} X_{it}^{1-\chi}$, where X_{it} is the amount of an intermediate input used by firm i. The intermediate input itself is produced from the same CES (43) aggregate as consumption. Observe that $1-\chi$ measures the extent of strategic complementarity in price-setting, since it makes firms' marginal cost more dependent on their competitors' prices. In appendix E.5, we show the following.

Proposition 5. *In the strategic complementarity model, proposition 4 continues to apply unchanged, except that the generalized Phillips curve (54) is now replaced by*

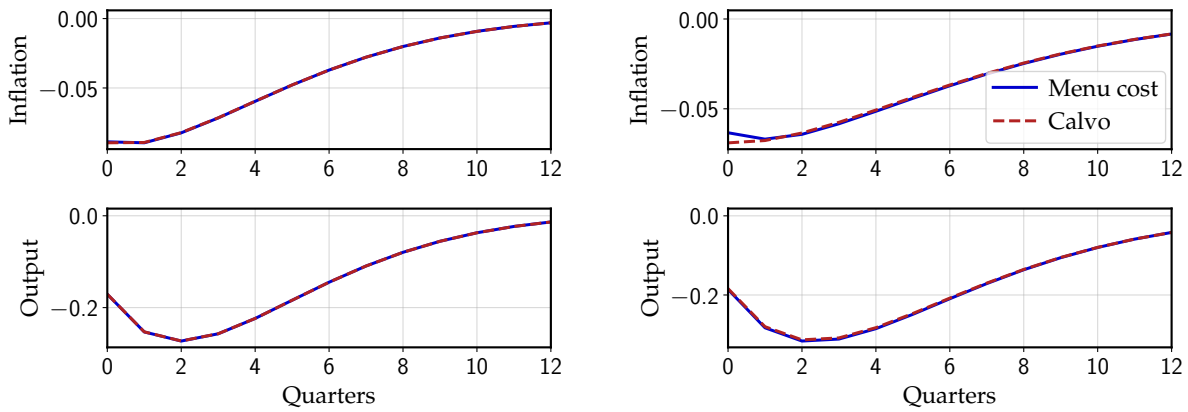
$$\boldsymbol{\pi} = (\varphi + \sigma)\, \chi \mathbf{K} \cdot \hat{\mathbf{Y}} \tag{56}$$

This generalizes an existing result from the Calvo literature that strategic complementarities scale down the slope parameter κ in the New Keynesian Phillips curve—exactly the same as one would obtain with less frequent price adjustment. Similarly, proposition 5 shows that in the menu cost model, strategic complementarity scales down the generalized Phillips curve $\mathbf{K}$ by χ. Since we have found that $\mathbf{K}$ is close to the Calvo NKPC, this amounts to scaling down κ, as in the standard result.

While the proof of proposition 5 requires the same formal $\sigma_\epsilon \to 0$ limit as in proposition 4, the basic logic only requires the concepts from section 2. With strategic complementarity, prices are given by $\hat{\mathbf{P}} = \Psi(\chi\widehat{\mathbf{MC}} + (1-\chi)\hat{\mathbf{P}})$, which can be rewritten in terms of real marginal cost as $\hat{\mathbf{P}} = \Psi\left(\chi\widehat{\mathbf{mc}} + \hat{\mathbf{P}}\right)$, where $\widehat{\mathbf{MC}}$ and $\widehat{\mathbf{mc}}$ are shocks to the marginal cost of labor. This is identical to (15), except with $\widehat{\mathbf{mc}}$ scaled down by χ. Hence the mapping $\mathbf{K}$ from $\widehat{\mathbf{mc}}$ and $\boldsymbol{\pi}$ derived in (16) is also scaled down by χ.

[43] This result provides an ex-post justification for a Krusell and Smith (1998) type algorithm used in the literature (e.g. Midrigan 2011) in which firms believe aggregates evolve a la Calvo.

[44] Alternative forms of strategic complementarities considered in the menu-cost literature include kinked demand curves (Klenow and Willis 2016) and oligopolistic competition (Mongey 2021). These "micro" strategic complementarities act differently from the "macro" strategic complementarity we consider here because they also narrow the *Ss* bands for given idiosyncratic shocks and menu costs. With an appropriate recalibration of the steady state, these can provide alternative microfoundations for our χ. Under a Calvo assumption, Gopinath and Itskhoki (2011) derive the implications of kinked demand curves for the aggregate Phillips curve, and Wang and Werning (2022) derive the more complex effects of oligopolistic competition.



(a) Golosov-Lucas (b) Nakamura-Steinsson

Figure 12: Impulse response to monetary shock for Smets-Wouters model with state-dependent pricing

Note: both pricing blocks were calibrated as in table 1. The approximating Calvo models are constructed in section (4). The other dynamics are those in Smets and Wouters (2007).

This result complements parallel work by Alvarez et al. (2022b), who characterize, in continuous-time models with strategic complementarity, the impulse response of prices $\hat{\mathbf{P}}$ to permanent *nominal* marginal cost shocks $\widehat{\mathbf{MC}}$. Because this impulse response satisfies $\hat{\mathbf{P}} = \chi \sum_{k\geq 0} \left((1-\chi)^k \Psi^{k+1} \right) \widehat{\mathbf{MC}}$, no simple scaling result comparable to (56) is attainable. Instead, their characterization exploits the fact that Ψ is self-adjoint and compact with a well-chosen inner product.[45]

5.3 Smets-Wouters model with menu costs

The logic behind proposition 4 continues to apply to a broader set of DSGE models: in the limit of small idiosyncratic shocks, the model with menu costs is equivalent to the model with Calvo pricing, but with the NKPC replaced by the generalized Phillips curve (54). To illustrate this result, figure 12 simulates a Smets and Wouters (2007) model with menu cost pricing. In this model, all equations and parameters are those from Smets and Wouters (2007) (estimated parameters are equal to their posterior means), with the exception of the price Phillips curve, which we replace by the generalized Phillips curve $\mathbf{K}$ of either the GL or the NS model, as well as by the respective approximating Calvo models. Now, the output responses to a monetary shock are more comparable across GL and NS, reflecting the presence of wage rigidities as an additional source of nominal rigidity. However, replacing the menu-cost model with its approximating Calvo model still provides an extremely close fit.

6 Obtaining the Generalized Phillips Curve from Micro Data

Our results so far have focused on the canonical menu cost model, with a two-point distribution of menu costs $\{0, \xi\}$. While this is a workhorse model in the literature, it has difficulty matching

[45] They call their continuous-time analog of Ψ the "kernel", since it is the kernel of an integral equation.

the empirical distribution of price changes. Instead, the data calls for a model with a generalized hazard function, as implied by a more general distribution of menu costs (Alvarez et al., 2022a).[46]

In this section, we extend our analysis to this case. We generalize our exact equivalence result to generalized hazards, and show that the pass-through matrix and the generalized Phillips curve of the model can be computed directly from the law of motion for price gaps, without the need to explicitly solve the optimizing model. Empirically, we show how to use the distribution of price changes, together with the frequency of adjustment, to infer this law of motion—allowing us to derive the full macro-level first-order relationship between real marginal cost and inflation from this micro data.

Allowing for a general distribution of menu costs. As in section 2.1, we consider a continuum of firms, each solving the cost minimization problem (2). Now, however, we assume that ξ_{it} is iid drawn from a general distribution with continuous cdf $\mathcal{H}(\cdot)$. The main implication of this change is that the law of motion is now no longer described by *Ss* bands $[\underline{x}_t, \overline{x}_t]$ and a reset price gap x_t^*; instead, there is a state-dependent generalized hazard function $\Lambda_t(x) \in [0,1]$ that captures the adjustment probability for a given price gap x at time t. The law of motion of price gaps is then given by

$$x_{it} = \begin{cases} x_t^* & \text{with probability } \Lambda_t(x_{it-1} - \epsilon_{it}) \\ x_{it-1} - \epsilon_{it} & \text{otherwise, with } \epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2) \end{cases} \tag{57}$$

Note that here we assume that ϵ_{it} is drawn from a normal distribution with variance σ_ϵ^2, as in our calibrations of the canonical menu cost model in section 2.5.

In the steady state, $\Lambda(x)$ is symmetric and $x^* = 0$. We continue to denote the stationary distribution of price gaps before adjustment by $g(x)$. The steady state distribution of price changes has the density $\Delta p \mapsto \Lambda(-\Delta p)g(-\Delta p)$. We continue to denote the frequency of adjustment by $\text{freq} = \int \Lambda(x)g(x)dx$, and expected price gaps by $E^t(x) \equiv \mathbb{E}[x_{it}|x_{i0} = x]$.

Generalizing proposition 1. In the exact equivalence result of section 3, only two TD models were necessary to describe the aggregate pricing behavior of the menu cost model. This is because there were only two margins of adjustment of policies in response to shocks: *Ss* bands could shift in parallel (the extensive margin), and the reset gap could shift (the intensive margin).

In the extended model, the entire hazard function shifts in response to shocks. Intuitively, there are more margins of adjustment, one for each level of the price gap x. Proposition 1 then has the following generalization.

Proposition 6. *The pass-through matrix of a generalized hazard model can be written as*

$$\Psi = \text{freq} \cdot \sum_{t=0}^{\infty} E^{t\prime}(0) \cdot \Psi^{\Phi^i} + \int \Lambda'(x)g(x) \cdot \left(\sum_{t=0}^{\infty} E^t(x) \right) \cdot \Psi^{\Phi^e(x)} dx \tag{58}$$

[46] See Berger and Vavra (2018), Caballero and Engel (1992, 1993), Gagnon, López-Salido and Vincent (2013), Luo and Villar (2021) for papers that estimate the empirical hazard function.

where $\Phi_t^i = E^{t\prime}(0)$ *as before; and* $\Phi_t^e(x) = E^t(x)/x$.

Here, the extensive margin in (58) depends on an entire integral of TD pass-through matrices with different survival functions $\Phi_t^e(x)$. Note that the weight for each x scales both with the density $g(x)$ and the derivative $\Lambda'(x)$ of the generalized hazard. The latter measures the sensitivity of the extensive margin at x: a high $\Lambda'(x)$ implies a locally very sensitive extensive margin that receives high weight, and vice versa.

To implement (58), all we need is the generalized hazard function Λ and the variance σ_ϵ^2 of idiosyncratic shocks. Together, these imply a stationary density $g(x)$ for price gaps, which can be calculated by iterating forward on the density; expected price gaps $E^t(x)$, which can be calculated by iteratively taking expectations; and the price reset frequency $\text{freq} = \int \Lambda(x)g(x)dx$.

In principle, with sufficiently rich panel data on price gaps, one could obtain the law of motion of price gaps directly from the data, and then implement equation (58) directly with this empirical law of motion. Given that, with some limited exceptions (e.g. Karadi, Schoenle and Wursten 2021), price gap data are difficult to come by, we will instead infer the law of motion from a more common type of data: the distribution and frequency of price changes.

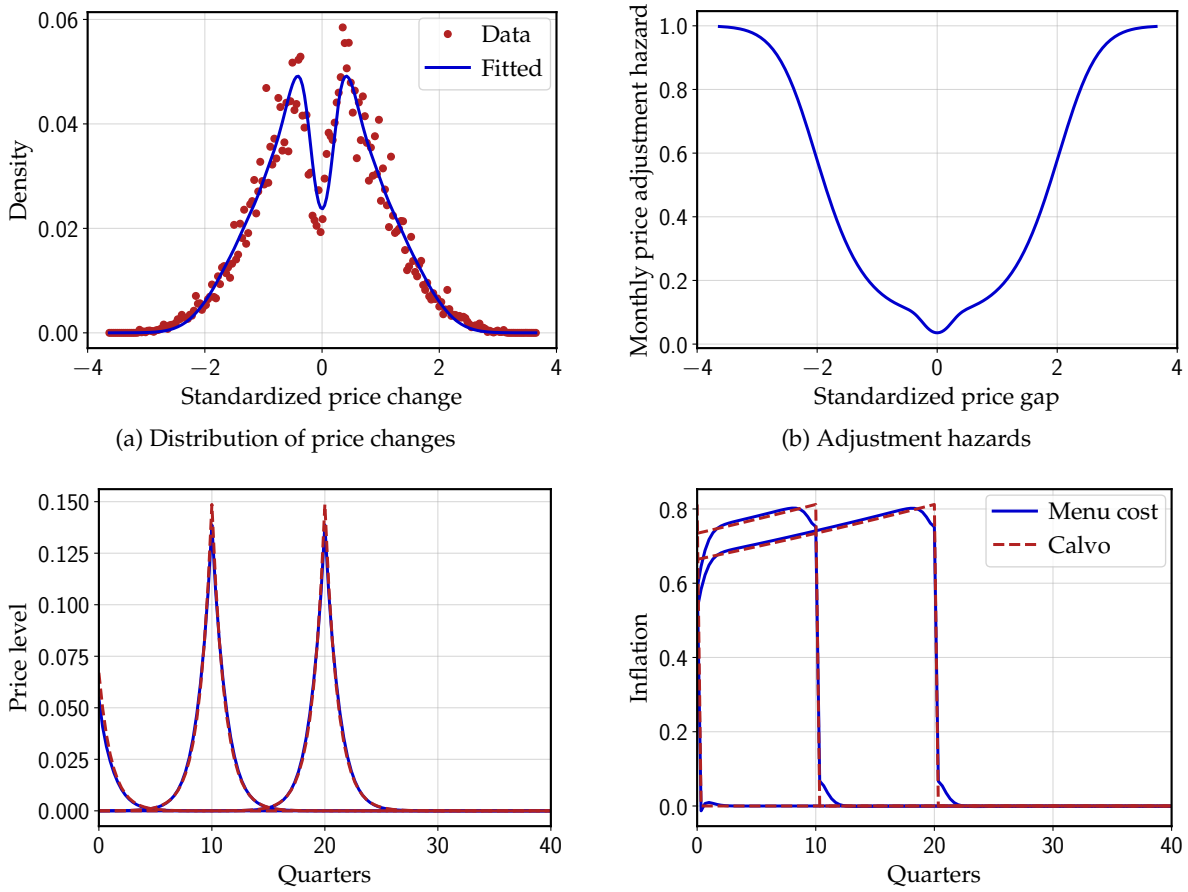
Obtaining Λ and σ_ϵ^2 from the data. Here, we briefly describe our procedure to back out the objects in (58) from price change data, and provide details in appendix F.2. We postulate a symmetric functional form, with several flexible parameters, for the generalized hazard Λ. Given guesses for these parameters and the variance σ_ϵ^2 of innovations, we iterate forward from an arbitrary initial density to obtain the steady-state density $g(x)$, using cubic splines to represent $g(x)$ and Gauss-Hermite quadrature to integrate over normal innovations $\mathcal{N}(0, \sigma_\epsilon^2)$. This process gives us a map from parameters to the steady-state density of price adjustments $\Lambda(\Delta p)g(\Delta p)$, and we solve for parameters that give the least-squares fit to the histogram of price changes in the price data (scaled by frequency). These parameters give us $\Lambda(x)$ and σ_ϵ^2, from which we evaluate (58) as discussed above.

Example from Israel. We show how this approach works for data on supermarket prices from Israel compiled by the Bank of Israel.[47] The top row in figure 13 shows the observed distribution of daily standardized price changes in the data (red) as well as the fitted curve. The hazard function is U shaped. The pass-through matrix and generalized Phillips curve are both well approximated by Calvo models, just as we had found for the canonical menu cost model.

7 Conclusion

In the past two decades, the growing availability of micro data on prices has spurred the development of a large literature that models price-setting decisions in presence of idiosyncratic shocks

[47] We thank Oleksiy Kryvtsov and Sigal Ribon for sharing the data with us.



(a) Distribution of price changes

(b) Adjustment hazards

(c) Pass-through matrix

(d) Generalized Phillips curve

Figure 13: Backing out the generalized Phillips curve from price change data alone

Note: The upper left panel shows the price change distribution from in data (red dots) and in the fitted model (blue line). The upper right panel shows the adjustment hazards that implement such a distribution (see appendix F.2). The bottom panels show the corresponding pass-through matrix, computed using 58, and the generalized Phillips curve.

and menu costs. In this paper, we show that these new models have the same first-order aggregate implications as older time-dependent models, provided that the hazard rates of price adjustment are suitably chosen. We provide sufficient statistic formulas to recover these virtual hazard rates—and therefore the generalized Phillips curve of the menu cost model—either directly from price change data, or from simulations of the steady state of the menu cost model.

We find that the generalized Phillips curve of these menu-cost models is very close to the Calvo Phillips curve, but with a higher slope. In our benchmark calibrations of the Golosov-Lucas and the Nakamura-Steinsson models, the slopes are $\kappa = 1.71$ and $\kappa = 0.47$, compared to a slope of $\kappa = 0.08$ in the Calvo model with the same frequency of price adjustment. By contrast, estimates based on macro data suggests that the slope of the aggregate Phillips curve may be even below this Calvo slope. For instance, Hazell, Herreño, Nakamura and Steinsson (2022) estimate $\kappa = 0.0031$, and point out that typical values in the macro literature are all below 0.05.[48]

[48]These numbers convert their Table III estimates by assuming that the elasticity of real marginal cost to unemploy-

We showed that strategic complementarities can, in principle, reconcile the micro and macro estimates: in our roundabout production model, they simply scale the generalized Phillips curve. It remains an open question, however, whether empirically plausible complementarities can lower the Phillips curve slope enough to match the recent macro estimates.

Simple strategic complementarities, however, cannot solve two broader issues with the New Keynesian Phillips curve: the lack of intrinsic inflation persistence (e.g. Fuhrer and Moore 1995, Galí et al. 2001), and the extreme forward-lookingness at the heart of the forward guidance puzzle (e.g. Del Negro et al. 2023). Multi-sector models with deviations from full information rational expectations and/or complex input-output linkages (e.g. Mankiw and Reis 2002, Woodford 2003a, Nimark 2008, Gabaix 2020, Angeletos and Huo 2021, Rubbo 2023), could be fruitfully combined with menu cost models to continue matching micro data on price changes while solving these broader issues that arise when confronting the generalized Phillips curve with the macro data.

Our work also leaves other open questions. For instance, our main results are only to first order. Although we found relatively small nonlinearities for aggregate shocks of typical magnitude in developed countries—generating up to a 5% increase in inflation—menu cost and Calvo models can still be quite different for larger shocks. Such large shocks are more common at the sectoral level, suggesting that menu costs should be studied in conjunction with models that disaggregate between sectors. Further, even when shocks are small, the second-order properties of menu cost vs. Calvo models will be crucial to understand their implications for optimal monetary policy. Finally, the first-order equivalence between state-dependent and time-dependent models may also find applications in other fields where fixed costs are an important component of decisions, such as those of adjusting capital (e.g. Khan and Thomas 2008) or purchasing a durable good (e.g. Berger and Vavra 2015).

ment is 2, as in the calibration of the textbook New Keynesian model in section 5.1 with $\varphi + \sigma = 2$.

References

Afrouzi, Hassan and Choongryul Yang, "Dynamic Rational Inattention and the Phillips Curve," Working Paper 3770462, Social Science Research Network, January 2021.

Alexandrov, Andrey, "The Effects of Trend Inflation on Aggregate Dynamics and Monetary Stabilization," *Manuscript*, June 2022.

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

_ **and** _ , "The Analytic Theory of a Monetary Shock," *Econometrica*, 2022, *90* (4), 1655–1680.

_ **and Pablo Andrés Neumeyer**, "The Passthrough of Large-Cost Shocks in an Inflationary Economy," in Gonzalo Castex, Jordi Galí, and Diego Saravia, eds., *Changing Inflation Dynamics, Evolving Monetary Policy*, Vol. 27 of *Series on Central Banking, Analysis, and Economic Policies*, Central Bank of Chile, 2020.

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

_ **,** _ **, and Juan Passadore**, "Are State- and Time-Dependent Models Really Different?," *NBER Macroeconomics Annual 2016*, May 2017, *31*, 379–457.

_ **,** _ **, and Takis Souganidis**, "Price Setting with Strategic Complementarities as a Mean Field Game," Working Paper 30193, National Bureau of Economic Research, July 2022.

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

_ **, Martin Beraja, Martín Gonzalez-Rozada, and Pablo Andrés Neumeyer**, "From Hyperinflation to Stable Prices: Argentina's Evidence on Menu Cost Models," *Quarterly Journal of Economics*, February 2019, *134* (1), 451–505.

Angeletos, George-Marios and Zhen Huo, "Myopia and Anchoring," *American Economic Review*, April 2021, *111* (4), 1166–1200.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, 2021, *89* (5), 2375–2408.

_ **, Matthew Rognlie, and Ludwig Straub**, "The Intertemporal Keynesian Cross," Working Paper 25020, National Bureau of Economic Research, May 2023.

Auer, Raphael, Ariel Burstein, and Sarah M. Lein, "Exchange Rates and Prices: Evidence from the 2015 Swiss Franc Appreciation," *American Economic Review*, February 2021, *111* (2), 652–686.

Bakhshi, Hasan, Hashmat Khan, and Barbara Rudolf, "The Phillips Curve Under State-Dependent Pricing," *Journal of Monetary Economics*, November 2007, *54* (8), 2321–2345.

Baley, Isaac and Andrés Blanco, "Aggregate Dynamics in Lumpy Economies," *Econometrica*, 2021, *89* (3), 1235–1264.

Ball, Laurence and N. Gregory Mankiw, "Asymmetric Price Adjustment and Economic Fluctuations," *The Economic Journal*, 1994, *104* (423), 247–261.

Barro, Robert J., "A Theory of Monopolistic Price Adjustment," *Review of Economic Studies*, 1972, *39* (1), 17–26.

Berger, David and Joseph Vavra, "Consumption Dynamics During Recessions," *Econometrica*, January 2015, *83* (1), 101–154.

_ **and** _ , "Dynamics of the U.S. Price Distribution," *European Economic Review*, April 2018, *103*, 60–82.

Bils, Mark and Peter J. Klenow, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, October 2004, *112* (5), 947–985.

Blanco, Andres, Corina Boar, Callum Jones, and Virgiliu Midrigan, "Nonlinear Inflation Dynamics in Menu Cost Economies," *Manuscript*, December 2022.

Bonadio, Barthélémy, Andreas M Fischer, and Philip Sauré, "The Speed of Exchange Rate Pass-Through," *Journal of the European Economic Association*, February 2020, *18* (1), 506–538.

Bonomo, Marco, Carlos Carvalho, Oleksiy Kryvtsov, Sigal Ribon, and Rodolfo Rigato, "Multiproduct Pricing: Theory and Evidence From Large Retailers," Working Paper 3590402, Social Science Research Network, May 2022.

Burstein, Ariel and Christian Hellwig, "Welfare Costs of Inflation in a Menu Cost Model," *American Economic Review*, May 2008, *98* (2), 438–443.

Caballero, Ricardo J. and Eduardo M. R. A. Engel, "Price Rigidities, Asymmetries, and Output Fluctuations," Working Paper 4091, National Bureau of Economic Research, June 1992.

_ **and** _ , "Microeconomic Rigidities and Aggregate Price Dynamics," *European Economic Review*, May 1993, *37* (4), 697–711.

_ **and** _ , "Price Stickiness in Ss Models: New Interpretations of Old Results," *Journal of Monetary Economics*, September 2007, *54, Supplement*, 100–121.

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

Caplin, Andrew and John Leahy, "State-Dependent Pricing and the Dynamics of Money and Output," *Quarterly Journal of Economics*, August 1991, *106* (3), 683–708.

_ and _ , "Aggregation and Optimization with State-Dependent Pricing," *Econometrica*, 1997, *65* (3), 601–625.

Carvalho, Carlos and Felipe Schwartzman, "Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models," *Journal of Monetary Economics*, November 2015, *76*, 141–156.

Cogley, Timothy and Argia M Sbordone, "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review*, November 2008, *98* (5), 2101–2126.

Coibion, Olivier and Yuriy Gorodnichenko, "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, February 2011, *101* (1), 341–370.

Costain, James and Anton Nakov, "Distributional Dynamics Under Smoothly State-Dependent Pricing," *Journal of Monetary Economics*, September 2011, *58* (6), 646–665.

Danziger, Leif, "A Dynamic Economy with Costly Price Adjustments," *American Economic Review*, September 1999, *89* (4), 878–901.

Dedola, Luca, Mark Strom Kristøffersen, and Gabriel Züllig, "The Extensive and Intensive Margin of Price Adjustment to Cost Shocks: Evidence from Danish Multiproduct Firms," *Manuscript*, April 2021.

Deimling, Klaus, *Nonlinear Functional Analysis*, Courier Corporation, 1985.

Dotsey, Michael, Robert G. King, and Alexander L. Wolman, "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics*, May 1999, *114* (2), 655–690.

Fernández-Villaverde, Jesús, "The Econometrics of DSGE Models," *SERIEs*, March 2010, *1* (1), 3–49.

_ , Juan Rubio-Ramírez, and Frank Schorfheide, "Chapter 9 - Solution and Estimation Methods for DSGE Models," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier, January 2016, pp. 527–724.

Fuhrer, Jeff and George Moore, "Inflation Persistence," *Quarterly Journal of Economics*, February 1995, *110* (1), 127–159.

Gabaix, Xavier, "A Behavioral New Keynesian Model," *American Economic Review*, August 2020, *110* (8), 2271–2327.

Gagnon, Etienne, "Price Setting during Low and High Inflation: Evidence from Mexico," *The Quarterly Journal of Economics*, August 2009, *124* (3), 1221–1263.

_ , David López-Salido, and Nicolas Vincent, "Individual Price Adjustment along the Extensive Margin," *NBER Macroeconomics Annual 2012*, May 2013, *27*, 235–281.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, February 2008.

_ and Mark Gertler, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, October 1999, *44* (2), 195–222.

_ , _ , and J. David López-Salido, "European Inflation Dynamics," *European Economic Review*, June 2001, *45* (7), 1237–1270.

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

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

Gopinath, Gita and Oleg Itskhoki, "In Search of Real Rigidities," *NBER Macroeconomics Annual 2010*, May 2011, *25* (1), 261–310.

Hall, Robert E., "A Major Shock Makes Prices More Flexible and May Result in a Burst of Inflation or Deflation," Working Paper 31025, National Bureau of Economic Research, March 2023.

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

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

_ , Raphael Schoenle, and Jesse Wursten, "Measuring Price Selection in Microdata: It's Not There," *Manuscript*, June 2021.

Kemeny, John G., J. Laurie Snell, and Anthony W. Knapp, *Denumerable Markov Chains*, Vol. 40 of *Graduate Texts in Mathematics*, New York, NY: Springer New York, 1976.

Khan, Aubhik and Julia K. Thomas, "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, March 2008, *76* (2), 395–436.

Klenow, Peter J. and Benjamin A. Malin, "Chapter 6 - Microeconomic Evidence on Price-Setting," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Elsevier, 2010, pp. 231–284.

_ **and Jonathan L. Willis**, "Real Rigidities and Nominal Price Changes," *Economica*, July 2016, *83* (331), 443–472.

_ **and Oleksiy Kryvtsov**, "State-Dependent or Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation?," *Quarterly Journal of Economics*, August 2008, *123* (3), 863–904.

Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896.

La'O, Jennifer and Alireza Tahbaz-Salehi, "Optimal Monetary Policy in Production Networks," *Econometrica*, 2022, *90* (3), 1295–1336.

Luo, Shaowen and Daniel Villar, "The Skewness of the Price Change Distribution: A New Touchstone for Sticky Price Models," *Journal of Money, Credit and Banking*, 2021, *53* (1), 41–72.

Maćkowiak, Bartosz and Mirko Wiederholt, "Business Cycle Dynamics under Rational Inattention," *Review of Economic Studies*, October 2015, *82* (4), 1502–1532.

Mankiw, N. Gregory and Ricardo Reis, "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, November 2002, *117* (4), 1295–1328.

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

Mongey, Simon, "Market Structure and Monetary Non-neutrality," Working Paper 29233, National Bureau of Economic Research, September 2021.

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

_ **and** _ , "Monetary Non-neutrality in a Multisector Menu Cost Model," *Quarterly Journal of Economics*, August 2010, *125* (3), 961–1013.

_ , _ **, Patrick Sun, and Daniel Villar**, "The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation," *Quarterly Journal of Economics*, November 2018, *133* (4), 1933–1980.

Negro, Marco Del, Marc Giannoni, and Christina Patterson, "The Forward Guidance Puzzle," *Journal of Political Economy Macroeconomics*, 2023, *forthcoming*.

Nimark, Kristoffer, "Dynamic Pricing and Imperfect Common Knowledge," *Journal of Monetary Economics*, March 2008, *55* (2), 365–382.

Reiter, Michael, "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, March 2009, *33* (3), 649–665.

Rubbo, Elisa, "Networks, Phillips Curves and Monetary Policy," *Econometrica*, 2023, *forthcoming*.

Rudin, Walter, *Principles of Mathematical Analysis*, 3 ed., McGraw-Hill, 1976.

Sbordone, Argia M., "Prices and Unit Labor Costs: A New Test of Price Stickiness," *Journal of Monetary Economics*, March 2002, *49* (2), 265–292.

Sheedy, Kevin D., "Intrinsic Inflation Persistence," *Journal of Monetary Economics*, November 2010, *57* (8), 1049–1061.

Sheshinski, Eytan and Yoram Weiss, "Inflation and Costs of Price Adjustment," *Review of Economic Studies*, June 1977, *44* (2), 287–303.

Simon, Herbert A., "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function," *Econometrica*, 1956, *24* (1), 74–81.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, June 2007, *97* (3), 586–606.

Taylor, John B., "Staggered Wage Setting in a Macro Model," *American Economic Review*, May 1979, *69* (2), 108–113.

_ , "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, February 1980, *88* (1), 1–23.

Tsiddon, Daniel, "The (Mis)Behaviour of the Aggregate Price Level," *Review of Economic Studies*, October 1993, *60* (4), 889–902.

Vavra, Joseph, "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation," *Quarterly Journal of Economics*, February 2014, *129* (1), 215–258.

Wang, Olivier and Iván Werning, "Dynamic Oligopoly and Price Stickiness," *American Economic Review*, August 2022, *112* (8), 2815–2849.

Whelan, Karl, "Staggered Price Contracts and Inflation Persistence: Some General Results," *International Economic Review*, 2007, *48* (1), 111–145.

Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," in Philippe Aghion, Roman Frydman, Joseph E. Stiglitz, and Michael Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton: Princeton University Press, 2003.

_ , *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, August 2003.

Yun, Tack, "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles," *Journal of Monetary Economics*, April 1996, *37* (2), 345–370.

Appendix for "New Pricing Models, Same Old Phillips Curves"

A Continuous-Time Version of Our Results

In this appendix, we extend our equivalence result between time- and state-dependent pricing models to a continuous time setting. We start by setting up a random menu cost model in continuous time, based on Alvarez et al. (2016) and Alvarez et al. (2022b), then move to time-dependent models, and finish by exploring the connection between both.

A.1 A random menu cost model

There is a continuum of firms, indexed by $i \in [0,1]$, each one selling a single product, whose at price at instant t is denoted p_{it}. Each firm has a static optimal price that is the sum of an idiosyncratic term p^*_{it} and a common nominal marginal cost component $MC(t)$. The idiosyncratic component is assumed to follow an i.i.d. Brownian motion without drift:

$$dp^*_{it} = -\sigma dW_{it}, \tag{A.1}$$

where W_{it} is a standard Wiener process. As in our discrete time setting, we work in a perfect-foresight environment, in which at $t = 0$ a path for $MC(t)$ is announced. Prior to $t = 0$, the economy is at steady state.

Firms may adjust their prices either by paying a fixed menu cost ξ or by receiving a random free adjustment opportunity, which arrives at a Poisson rate $\lambda \in [0, \infty)$. At any given instant, firms face economic losses for not charging their optimal prices. As in the discrete time case, define the price gap $x_{it} = p_{it} - p^*_{it}$. Losses are then given by $\frac{1}{2}(x_{it} - MC(t))^2$.

In the absence of price adjustments, x_{it} follows the Brownian motion

$$dx_{it} = \sigma dW_{it}.$$

Moreover, it is well known that the presence of a fixed menu cost generates an optimal policy that takes the form of an inaction region a reset point. Since the common shock $MC(t)$ evolves over time, the optimal policy is also time-varying, with the inaction interval taking the form $(\underline{x}(t), \overline{x}(t))$ and the reset point denoted $x^*(t)$.

We can state the firm problem recursively in terms of the state variable x, as follows. There is a value function $V(t,x)$, which obeys the following Hamilton-Jacobi-Bellman (HJB) equation inside the inaction region:

$$\rho V(t,x) = \frac{1}{2}\left(x - MC(t)\right)^2 + \frac{\sigma^2}{2}\partial_{xx}V(t,x) + \lambda\left(V(t,x^*(t)) - V(t,x)\right) + \partial_t V(t,x), \text{ for } x \in (\underline{x}(t), \overline{x}(t)), \tag{A.2}$$

where ρ is the discount rate. At the boundaries of the inaction region, we have the following value matching and smooth pasting conditions, which, together with the optimality condition for the reset point $x^*(t)$, complete the recursive characterization of the problem:

$$V(t, \underline{x}(t)) = V(t, \overline{x}(t)) = V(t, x^*(t)) + \xi,$$

$$\partial_x V(t, \underline{x}(t)) = \partial_x V(t, \overline{x}(t)) = \partial_x V(t, x^*(t)) = 0.$$

Given the optimal policies in response to the shock $MC(t)$, the next step is to compute the distribution of price gaps. Let $g(t, x)$ be the probability density function of price gaps at time t. It evolves according to a Kolmogorov forward equation:

$$\partial_t g(t, x) = \frac{\sigma^2}{2} \partial_{xx} g(t, x), \tag{A.3}$$

equipped with the following conditions:

$$g(t, \underline{x}(t)) = g(t, \overline{x}(t)) = 0,$$

$$g(t, x) \text{ continuous at } x^*(t),$$

$$\int_{\underline{x}(t)}^{\overline{x}(t)} g(t, x)\, dx = 1.$$

This equation is solved forward, with the initial condition at $t = 0$ being steady state distribution. As in the discrete time case, deviations of the price level to its steady state values, in logs, are given by

$$p(t) = \int x\, g(t, x)\, dx. \tag{A.4}$$

By first applying the HJB equation (A.2), followed by the KFE (A.3), one can compute the price level response to any nominal marginal cost shock.

A.2 Time-dependent models

As in discrete time, price setting in a time-dependent model is governed by a survival function $\Phi(s)$. Prices are randomly selected to adjust depending only on the time elapsed since the last adjustment, and $\Phi(s)$ is the probability that a price remains fixed for a time interval of length $\geq s$. This immediately implies $\Phi(0) = 1$. Again, each firm has a price gap $x_{it} = p_{it} - p_{it}^*$, with p_{it}^* evolving as in (A.1). Upon adjustment, firms choose a reset gap that solves

$$x^*(t) = \arg\max_x \frac{1}{2} \mathbb{E}_t \int_t^\infty e^{-\rho(s-t)} \Phi(s-t) \left(x + p_{it}^* - p_{is}^* - MC(s)\right)^2 ds,$$

which is given by

$$x^*(t) = \frac{1}{\int_0^\infty e^{-\rho s}\Phi(s)\,ds} \int_0^\infty e^{-\rho s}\Phi(s)MC(t+s)\,ds. \tag{A.5}$$

One can compute the log of the aggregate price level from past pricing decisions as

$$p(t) = \frac{1}{\int_0^\infty \Phi(s)\,ds} \int_0^\infty \Phi(s)\,x^*(t-s)ds. \tag{A.6}$$

There is a clear analogy between the above expressions and their discrete time counterparts from section 2.

A.3 The pass-through operator

Both classes of models above generate a mapping from the aggregate marginal cost path $MC(t)$ to the price level $p(t)$, which we denote

$$p(t) = \mathcal{P}(t; \{MC(s)\}).$$

This can be linearized around $MC(t) = 0$ to obtain the first-order impulse response

$$p(t) = \int_0^\infty \Psi(t,s)MC(s)\,ds. \tag{A.7}$$

We call the operator on the right hand side of (A.7) the *pass-through operator*. (Alvarez et al. (2022b) call it the *kernel.)* Similarly to the discrete time case, for time-dependent models it is given by the composition of the operators in (A.5) and (A.6).

Before proceeding to the exact equivalence result in continuous time, it is necessary to generalize the notion of a survival function. From the definition of a survival function, it is clear that $\Phi(0) = 1$. Expressions (A.5) and (A.6), however, still define mathematically consistent mappings when $\Phi(s) \neq 1$ and even in cases where $\Phi(s) \to \infty$ as $s \to 0$, as long as this function has a finite integral on $[0, \infty)$.[A-1] When working with time-dependent pass-through operators, we allow for this possibility and refer to Φ as a *generalized survival function*.

A.4 Exact equivalence in continuous time

Consider a random menu cost model in steady state. Given the symmetry of the problem, the inaction region can be written as $(-\overline{x}, \overline{x})$ and the reset point is $x^* = 0$. Let x_t be the price gap of a

[A-1] Of course, it must still satisfy the other required properties of a survival function: it must be non-increasing, converge to zero as $s \to \infty$, and start at a positive (but not necessarily finite) $\Phi(0) > 0$.

firm that follows this optimal policy and define

$$E(t,x) = \mathbb{E}\left[x_t | x_0 = 0\right].$$

Exactly as in discrete time, the pass-through operator of the state-dependent model Ψ can be written as

$$\Psi = \alpha \Psi^{\Phi^e} + (1-\alpha)\,\Psi^{\Phi^i}.$$

The intensive margin pass-through operator Ψ^{Φ^i} is associated with the following generalized survival function:

$$\Phi^i(t) = \partial_x E(t,0),$$

while extensive margin component arises from the generalized survival function

$$\Phi^e(t) = \partial_x E(t,\overline{x}).$$

As we explain in more detail below, even though we call Φ^i a generalized survival function, it satisfies $\Phi^i(0) < \infty$ and could therefore be normalized to become a proper survival function. On the other hand, we have $\Phi^e(0) = \infty$, and so we must interpret the extensive margin component in the generalized sense. The weight α is given by

$$\alpha = f \times \int_0^\infty \partial_x E(t,0)\, dt,$$

where f is the flow of price adjustments, i.e.,

$$f = \lim_{\Delta t \to 0} \frac{\text{fraction of prices that change in } (t, t+\Delta t)}{\Delta t}.$$

The advantage of the continuous time approach is that it is possible to solve for $E(t,x)$ explicitly. First, it satisfies the Kolmogorov backward equation:

$$\frac{\partial E}{\partial t}(t,x) = \frac{\sigma^2}{2}\frac{\partial^2 E}{\partial x^2}(t,x) - \lambda E(t,x), \tag{A.8}$$

$$E(t,\underline{x}) = E(t,\overline{x}) = 0 \text{ for } t > 0,$$

$$E(0,x) = x.$$

Alvarez and Lippi (2022) provide a closed-form solution to this equation, which we reproduce here. Define

$$\eta_j = -\left[\lambda + \frac{\sigma^2}{2}\left(\frac{j\pi}{2\overline{x}}\right)^2\right],$$

$$\varphi_j(x) = \frac{1}{\sqrt{\overline{x}}}\sin\left(\frac{x+\overline{x}}{2\overline{x}} j\pi\right),$$

for $j = 1, 2, 3, \ldots$. These are the eigenvalues and corresponding eigenfunctions of the operator on the right hand side of (A.8). Moreover, let

$$b_j = \frac{4\overline{x}^{3/2}}{j\pi},$$

which are the coefficients one obtains from projecting the function $f(x) = x$ onto the eigenfunctions above.

Having defined these objects, we can express $E(t, x)$ as

$$E(t, x) = \sum_{j \text{ even}} e^{\lambda_j t} b_j \varphi_j(x).$$

Only even terms appear in the above summation because the function $f(x) = x$ is odd, and is therefore orthogonal to the eigenfunctions with odd indices, which are even functions.

From the previous result, we can compute $\Phi^i(t)$ as

$$\partial_x E(t, 0) = 2 \sum_{j \text{ even}} (-1)^{j/2} e^{\eta_j t}$$

and $\Phi^e(t)$ as

$$\partial_x E(t, \overline{x}) = 2 \sum_{j \text{ even}} e^{\eta_j t}.$$

Note that it immediately follows that $\Phi^e(0) = \infty$. We can also compute the associated adjustment hazards, defined as

$$\lambda^e(t) = -\frac{\partial}{\partial t} \log \Phi^e(t) = -\frac{\sum_{j \text{ even}} \eta_j e^{\eta_j t}}{\sum_{j \text{ even}} e^{\eta_j t}},$$

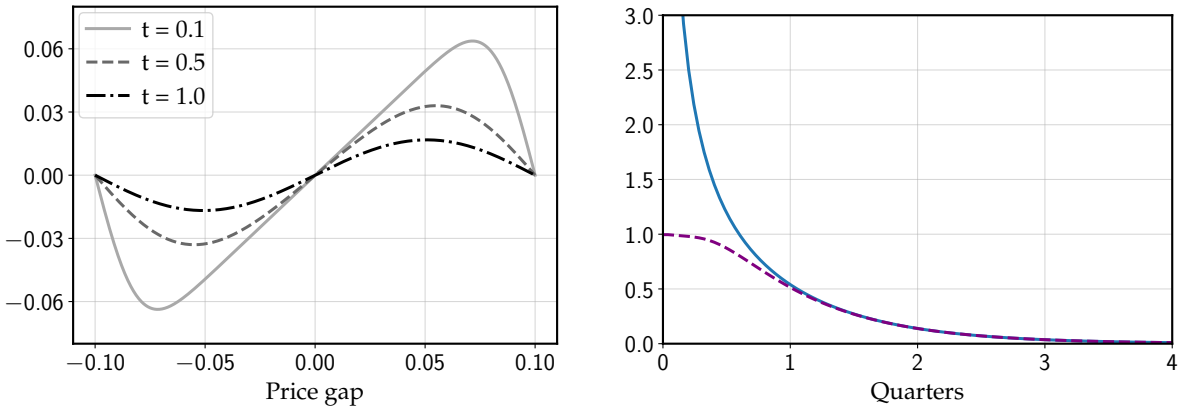
$$\lambda^i(t) = -\frac{\partial}{\partial t} \log \Phi^i(t) = -\frac{\sum_{j \text{ even}} (-1)^{j/2} \eta_j e^{\eta_j t}}{\sum_{j \text{ even}} (-1)^{j/2} e^{\eta_j t}}.$$

From the expressions above, it follows immediately that

$$\lim_{t \to \infty} \lambda^e(t) = \lim_{t \to \infty} \lambda^i(t) = -\lambda_2,$$

echoing our analogous result for the discrete time case (proposition 2). Figure A.1 shows $E(t, x)$, the generalized survival functions, and the corresponding adjustment hazards for illustrative parameter values. Notice that the extensive margin adjustment hazard $\lambda^e(t)$ must also be interpreted in a generalized sense, since $\lim_{t \to 0} \lambda^e(t) = \infty$.[A-2]

[A-2] We conjecture that $\lambda^e(t)$ can also be interpreted as the arrival hazard of a point process with countably many points but infinite average density.



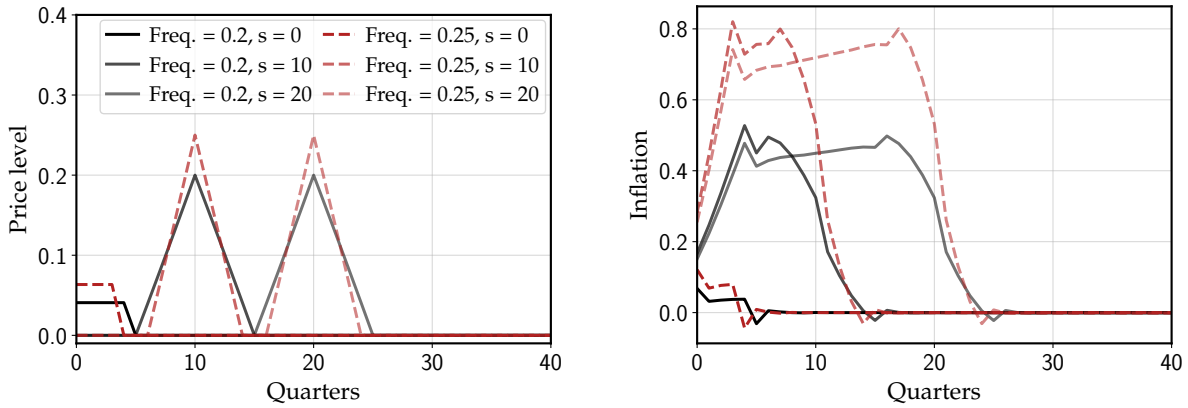
(a) $E(t,x)$

(b) $\Phi^e(t)$ and $\Phi^i(t)$

(c) $\lambda^e(t)$ and $\lambda^i(t)$

Figure A.1: Expected price gaps and generalized survival functions and hazards

Note: illustrative calibration with parameter values $\sigma = 0.05$, $\lambda = 0.1$, and the menu cost is such that $\overline{x} = 0.1$.



(a) Pass-through matrix

(b) Generalized Phillips curve

Figure B.1: Columns $s \in \{0, 10, 20\}$ of the Taylor model pass-through and generalized Phillips curve matrices

Note: Parameter values are $\tau \in \{4, 5\}$, corresponding to adjustment frequencies 0.2 and 0.25. The "wild" shape in Panel b is exact and does not contain numerical error.

B Appendix to Section 2

B.1 The Taylor model

Here, we show the pass-through matrix and generalized Phillips curve of a Taylor model. The Taylor model is characterized by a degenerate distribution of price spells: there is a parameter τ such that firms adjust with certainty every τ periods. One disadvantage of this model is that the frequency of adjustment is $1/\tau$, and τ must be an integer, so this model is not able to generate an arbitrary adjustment frequency in $[0, 1]$, as the Calvo model, for example.

Figure B.1 shows results. The left panel shows the pass-through matrix. As no price lasts longer than τ periods without being adjusted, a one-time nominal marginal cost shocks at date s only affects the price level in dates $s - \tau$ through $s + \tau$, so most entries of the matrix are exactly zero. The right panel shows the generalized Phillips curve. The discontinuous appearance of its columns come from the "discontinuity" in adjustment hazards: in the Taylor model, adjustment hazards are zero and then suddenly jump to 1 after τ periods.

B.2 Characterizing steady-state policy and distribution

We can rewrite the steady-state version of (2) recursively with the Bellman equation

$$V^n(x) \equiv \frac{1}{2}x^2 + \beta(1-\lambda)\mathbb{E}\left[\min(V^{n-1}(x+\epsilon), \xi + \min_{x^*} V^{n-1}(x^*))\right] + \beta\lambda \min_{x^*} V^{n-1}(x^*) \qquad \text{(A.9)}$$

whose fixed point is the value function $V(x)$ given a post-adjustment price gap of x (not including any costs already paid to adjust).

Reducing to bounded V on an interval $[-M, M]$. First, we observe that any value function V should satisfy $V(x) \geq \frac{1}{2}x^2$. Second, we have $V(0) \leq \frac{\beta}{1-\beta}(1-\lambda)\xi$, where the right is the value from the feasible policy of always adjusting to stay at $x = 0$.

It follows that for sufficiently large x (i.e. $|x| \geq M$ for some M), $V(x)$ must be strictly greater than $\xi + V(0)$. Hence, for these x the price-setter strictly prefers to adjust, and we have $\min(V(x), \xi + \min_{x^*} V(x^*)) = \min_{x^*} V(x^*)$. It is therefore not necessary to keep track of V outside $[-M, M]$ to evaluate (A.9) inside $[-M, M]$, or to obtain the optimal policy anywhere. Hence, when analyzing (A.9), we restrict ourselves to $[-M, M]$. Further, we note that the value function satisfies $V(x) \leq \frac{1}{2}M^2 + \frac{\beta}{1-\beta}(1-\lambda)\xi$ for all $x \in [-M, M]$, so that we can restrict our attention to bounded V.

From now on, our restriction to $[-M, M]$ will be implicit. We will assume that M is picked to be large enough that, for all parameters we consider, the firm always adjusts for $|x| \geq m$, for some $m < M$.

Characterizing value function steps. Suppose that $V^{n-1}(x)$ is nonnegative, symmetric around 0, continuously differentiable, and satisfies $(V^{n-1})'(x) > 0$ for $x > 0$. Suppose also that it satisfies $V^{n-1}(x) \geq \frac{1}{2}x^2$ and $V^{n-1}(0) \leq \frac{\beta}{1-\beta}(1-\lambda)\xi$.

It follows that the minimum will be at $x^* = 0$, and also that there will be some $0 < \bar{x} < M$[A-3] such that $V^{n-1}(x) < \xi + V^{n-1}(0)$ for all $x > \bar{x}$, and symmetrically for $x < -\bar{x}$. This allows us to replace (A.9) by the more specific

$$V^n(x) = \frac{1}{2}x^2 + \beta(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x'-x)V^{n-1}(x')dx' + \beta(1-\lambda)(V^{n-1}(0)+\xi)\left(1 - \int_{-\bar{x}}^{\bar{x}} f(x'-x)dx'\right) + \beta\lambda V^{n-1}(0) \quad \text{(A.10)}$$

where $\bar{x}$ is implicitly determined by the equation $V^{n-1}(\bar{x}) = \xi + V^{n-1}(0)$. It follows directly from (A.10) and the symmetry of f that V^n will also be nonnegative, symmetric around 0, and satisfy $V^n(x) \geq \frac{1}{2}x^2$ and $V^n(0) \leq \frac{\beta}{1-\beta}(1-\lambda)\xi$. All that remains is investigate the derivative.

Substituting $u = x' - x$, (A.10) can be rewritten as

$$V^n(x) = \frac{1}{2}x^2 + \beta(1-\lambda)\int_{-\bar{x}-x}^{\bar{x}-x} f(u)V^{n-1}(x+u)du + \beta(1-\lambda)(V^{n-1}(0)+\xi)\left(1 - \int_{-\bar{x}-x}^{\bar{x}-x} f(u)du\right) + \beta\lambda V^{n-1}(0)$$

[A-3] $\bar{x} < M$ follows from the choice of M above, while $0 < \bar{x}$ follows from the continuity of $V^{n-1}(x)$, since for x close enough to 0, $V^{n-1}(x)$ gets arbitrarily close to $V^{n-1}(0)$ and therefore below $\xi + V^{n-1}(0)$.

Now, differentiating with respect to x using Leibniz's rule, we find

$$\begin{aligned}(V^n)'(x) &= x + \beta(1-\lambda)\left(-f(\bar{x}-x)V^{n-1}(\bar{x}) + f(-\bar{x}-x)V^{n-1}(-\bar{x}) + \int_{-\bar{x}-x}^{\bar{x}-x}(V^{n-1})'(x+u)du\right) \\ &\quad - \beta(1-\lambda)(\xi + V^{n-1}(0))\left(-f(\bar{x}-x)V^{n-1}(\bar{x}) + f(-\bar{x}-x)V^{n-1}(-\bar{x})\right) \\ &= x + \beta(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x'-x)(V^{n-1})'(x')dx' \end{aligned} \tag{A.11}$$

where the cancellation follows from $V^{n-1}(\bar{x}) = V^{n-1}(-\bar{x}) = \xi + V^{n-1}(0)$. It follows that V^n is continuously differentiable.

Now, note that using the symmetry of f, we can rewrite (A.11) as

$$(V^n)'(x) = x + \beta(1-\lambda)\int_0^{\bar{x}}(f(x'-x) - f(-x'-x))(V^{n-1})'(x')dx' \tag{A.12}$$

The single-peakedness of f implies that $f(x'-x) - f(-x'-x) > 0$ for all $x, x' > 0$, so $(V^n)'(x) > 0$ for $x > 0$ follows from $(V^{n-1})'(x') > 0$ for $x' > 0$.

Explicitly constructing V through value function iteration and obtaining properties of the optimal policy. Write $V^0(x) \equiv \frac{1}{2}x^2$, which satisfies all hypotheses put on V^{n-1} in the previous discussion, and construct the series $\{V^n\}$ recursively. By induction, each V^n must be nonnegative, symmetric around 0, continuously differentiable, satisfy $(V^n)'(x) > 0$ for $x > 0$, and satisfy $V^n(x) \geq \frac{1}{2}x^2$ and $V^n(0) \leq \frac{\beta}{1-\beta}(1-\lambda)\xi$.

Now, by standard arguments, the right side of (A.9) is a contraction (in the sup norm) of modulus β. Hence the V^n converge uniformly to some fixed point V. This V must be nonnegative, symmetric around 0, weakly increasing for $x > 0$, and satisfy $V(x) \geq \frac{1}{2}x^2$ and $V(0) \leq \frac{\beta}{1-\beta}(1-\lambda)\xi$. Following the same logic as before, the set of points $x \geq 0$ for which $V(x) = \xi + V(0)$ must a closed subset of $(0, M)$. Let the minimum of this set be $\bar{x}$; then (A.10) holds with this $\bar{x}$.

Now, directly differentiating (A.10), the continuous differentiability of V follows from that of f. Hence (A.12) holds as a fixed point for V as well.

V is weakly increasing and must increase by at least ξ from $V(0)$ to $V(\bar{x})$, so we have $V'(x) \geq 0$ everywhere with $x > 0$ and $V'(x) > 0$ for some subset of $x > 0$ of positive measure. It then follows from (A.12) that since $f(x'-x) - f(-x'-x) > 0$ for all $x, x' > 0$, we have $V'(x) > 0$ for all $x > 0$ (and similarly $V'(x) < 0$ for $x < 0$ and $V'(0) = 0$).

This implies that there is a unique $\bar{x}$ satisfying $V(\bar{x}) = \xi + V(0)$. Hence, we have derived an optimal Ss policy, where there is no adjustment when x is in the interval $[-\bar{x}, \bar{x}]$, and always adjustment to 0 outside of this interval.

Finally, we note that differentiating (A.11) around the fixed point, we get

$$V''(x) = x + \beta(1-\lambda)\int_0^{\bar{x}}(f'(-x'-x) - f'(x'-x))V'(x')dx' \tag{A.13}$$

so that the continuous differentiability of V' follows from that of f. For the special case $x = 0$, we note that by assumption $f'(-x') = -f'(x') > 0$ for all $x' > 0$, so that the integral in (A.13) is positive and $V''(0) > 0$.

Steady-state distribution. We have shown above that in the steady-state version of the pricing problem, firms follow an *Ss* policy with adjustment to 0 whenever the price gap is outside the interval $[-\bar{x}, \bar{x}]$ (or there is a free adjustment with probability λ).

This implies a law of motion $\mathcal{T}$ for the density g *prior to adjustment* given by

$$(\mathcal{T}g)(x') = \int_{-\infty}^{\infty} p(x, x')g(x)dx \tag{A.14}$$

where the transition density $p(x, x')$ from x to x', satisfying $\int p(x, x')dx' = 1$, is given by

$$p(x, x') \equiv \begin{cases} f(x' - x) & |x| \leq \bar{x} \\ f(x') & |x| > \bar{x} \end{cases} \tag{A.15}$$

Defining $v(x') \equiv \min_{|x| \leq \bar{x}} f(x' - x)$, we note that by our assumptions on f that $v(x') > 0$ for all x'. Defining $h(x') \equiv \min(f(x'), v(x'))$, we have $p(x, x') \geq h(x') > 0$ for all x, x', and we can rewrite (A.14) for any density g with integral 1 as

$$(\mathcal{T}\pi)(x') = h(x') + \int_{-\infty}^{\infty} \left(p(x, x') - h(x')\right) g(x)dx$$

where $p(x, x') - h(x') \geq 0$. It follows that for any two densities g_1 and g_2 that we have

$$\begin{aligned}
\int_{-\infty}^{\infty} |(\mathcal{T}(g_1 - g_2))(x')|dx' &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \left(p(x, x') - h(x')\right) \left(g_1(x) - g_2(x)\right)dx \right| dx' \\
&\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(p(x, x') - h(x')\right) |g_1(x) - g_2(x)|dxdx' \\
&= \left(\int_{-\infty}^{\infty} |g_1(x) - g_2(x)| \left(\int_{-\infty}^{\infty} \left(p(x, x') - h(x')\right) dx' \right) dx \right) \\
&= \left(1 - \int_{-\infty}^{\infty} h(x')dx' \right) \int_{-\infty}^{\infty} |g_1(x) - g_2(x)|dx
\end{aligned}$$

so that, defining $\|g\| \equiv \int_{-\infty}^{\infty} |g(x)|dx$ to be the L^1 norm and $\mathcal{H} \equiv \int_{-\infty}^{\infty} h(x')dx' > 0$, we have

$$\|\mathcal{T}(\pi_1 - \pi_2)\| \leq (1 - \mathcal{H})\|\pi_1 - \pi_2\| \tag{A.16}$$

implying that $\mathcal{T}$ is a contraction on densities in the L^1 norm with modulus $1 - \mathcal{H}$.

It follows that there is a unique stationary density g that is a fixed point of the contraction $\mathcal{T}$, and that we will reach this density by starting with any $g^0 \in L^1$ and repeatedly iterating. Since $\mathcal{T}$ with symmetric f preserves symmetry around 0, it follows that the stationary g must be symmetric as well (since we can start with symmetric g^0 and iterate). It also follows from (A.14)–(A.15) and

the continuous differentiability of f that g is continuously differentiable.

B.3 Envelope result and contraction

In this section, we obtain some useful further technical results for the canonical menu cost model. In particular, we show that the backward mapping on V is differentiable, and indeed a contraction in a certain norm that regulates both the level and derivative of V. Hence, iterating backward in response to any sequence of first-order aggregate cost shocks, we retain Ss policies $(\underline{x}, \bar{x}, x^*)$, which are differentiable with respect to the shocks.

To start, consider the space of value functions V on $[-M, M]$ that are bounded and have bounded first derivative, endowed with the norm

$$\|V\| \equiv \sup_x |V(x)| + \zeta \sup_x |V'(x)| \tag{A.17}$$

for some $\zeta > 0$. Note that this space is complete (a Banach space).[A-4] Note also that in this norm, in a neighborhood around the steady-state V derived in the previous section, the adjustment policy will still be Ss. Indeed, if we write the mapping $T : (V_+, c) \to V$ from V_+ in this space and a cost scalar c to V also in this space[A-5], given locally by the Bellman equation

$$\begin{aligned} V(x) = \frac{1}{2}(x-c)^2 + \min_{x^*, \underline{x}, \bar{x}} \Big[& \beta(1-\lambda) \int_{\underline{x}}^{\bar{x}} f(x'-x) V_+(x') dx' \\ & + \beta(1-\lambda)(V_+(x^*) + \xi) \left(1 - \int_{\underline{x}}^{\bar{x}} f(x'-x) dx'\right) + \beta\lambda V_+(x^*) \Big] \end{aligned} \tag{A.18}$$

then the optimum is still characterized by the value matching conditions $V_+(\bar{x}) + \zeta = V_+(0)$ and $V_+(\underline{x}) + \zeta = V_+(0)$ and the first-order condition $V_+'(x^*) = 0$; and from the implicit theorem, these optima are differentiable with respect to V_+ in this norm around the steady state, with $d\bar{x} = -dV_+(\bar{x})/V_+'(\bar{x})$, $d\underline{x} = -dV_+(-\bar{x})/V_+'(-\bar{x})$, and $dx^* = -dV_+'(0)/V_+''(0)$.

Since the policy is differentiable with respect to V_+, we have a simple envelope result (obtainable simply by differentiating (A.18)), where in response to a perturbation dV_+, (A.18) becomes

$$dV(x) = \beta(1-\lambda) \int_{-\bar{x}}^{\bar{x}} f(x'-x) dV_+(x') dx' + \beta(1-\lambda) \left(1 - \int_{\underline{x}}^{\bar{x}} f(x'-x) dx'\right) dV_+(0) + \beta\lambda dV_+(0) \tag{A.19}$$

Note that this is a bounded map from dV_+ to dV in our normed space. First, it is immediate from (A.19) that $\sup_x |dV(x)| \leq \beta \sup_x |dV_+(x)|$.

[A-4] To see this, start by noting that for any Cauchy sequence $\{V_n\}$ in this norm, $\{V_n\}$ and $\{V_n'\}$ will also be Cauchy sequences in the ordinary sup norm, and therefore both individually converge to some limits. Then, the only remaining question to determine whether $\{V_n\}$ converges in our norm is whether the limit of V_n' equals the derivative of the limit of V_n; this, in turn, is a standard result in real analysis when there is uniform convergence (see e.g. Rudin (1976), Theorem 7.17).

[A-5] Boundedness of V follows immediately from the Bellman equation, and of of V' follows from differentiating as in (A.11).

Second, if we differentiate (A.19), we obtain

$$dV'(x) = \beta(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x'-x)dV'_+(x')dx' - \beta(1-\lambda)f(\bar{x}-x)\left(dV_+(\bar{x}) - dV_+(0)\right) + \beta(1-\lambda)f(-\bar{x}-x)\left(dV_+(-\bar{x}) - dV_+(0)\right)$$

and hence

$$\sup_x |dV'(x)| \leq \beta(1-\lambda)\sup_x |dV'_+(x)| + \beta(1-\lambda)\left(\sup_x f(\bar{x}-x) + f(-\bar{x}-x)\right) 4\sup_x |dV_+(x)|$$

and if we define the weight ζ in our norm (A.17) to be $\zeta \equiv \frac{(1-\beta)/2}{4\beta(1-\lambda)(\sup_x f(\bar{x}-x)+f(-\bar{x}-x))}$, this reduces to just

$$\sup_x |dV'(x)| \leq \beta(1-\lambda)\sup_x |dV'_+(x)| + \frac{1-\beta}{2}\zeta^{-1}\sup_x |dV_+(x)|$$

and we have

$$\begin{aligned}
\|V\| &= \sup_x |dV(x)| + \zeta \sup_x |dV'(x)| \\
&\leq \beta \sup_x |dV_+(x)| + \beta(1-\lambda)\zeta \sup_x |dV'_+(x)| + \frac{1-\beta}{2}\sup_x |dV_+(x)| \\
&< \frac{1+\beta}{2}\left(\sup_x |dV_+(x)| + \zeta \sup_x |dV'_+(x)|\right) \\
&= \frac{1+\beta}{2}\|V_+\|
\end{aligned}$$

and we conclude that the derivative mapping dV_+ to dV is a contraction with modulus $\frac{1+\beta}{2}$ in our norm.

B.4 Proof of the ξ/σ_ϵ^2 scaling property

In this section we prove that the first-order price level response $\{\log P_t\}$ to an arbitrary marginal cost shock $\{\log MC_t\}$ remains unchanged if σ_ϵ^2 and ξ are scaled by the same factor.

Scaling σ_ϵ^2 and ξ by factor $\vartheta^2 > 0$, the profit maximization problem (2) becomes

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2}\left(x_{it} - \log MC_t\right)^2 + \vartheta^2 \xi_{it} 1_{\{x_{it} \neq x_{it-1} - \vartheta\epsilon_{it}\}}\right]$$

Dividing the objective by ϑ^2, this is equivalent to

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2}\left(\vartheta^{-1}x_{it} - \vartheta^{-1}\log MC_t\right)^2 + \xi_{it} 1_{\{\vartheta^{-1}x_{it} \neq \vartheta^{-1}x_{it-1} - \epsilon_{it}\}}\right]$$

Denoting $\vartheta^{-1}x_{it}$ by $\tilde{x}_{it}$, this becomes

$$\min_{\{\tilde{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} \left(\tilde{x}_{it} - \vartheta^{-1} \log MC_t \right)^2 + \xi_{it} 1_{\{\tilde{x}_{it} \neq \tilde{x}_{it-1} - \epsilon_{it}\}} \right] \tag{A.20}$$

where the aggregate price level, stated in terms of $\tilde{x}_{it}$, is given by

$$\log P_t = \vartheta \int \tilde{x}_{it} di$$

Note that (A.20) has exactly the same structure as (2). Thus, if we denote the mapping from nominal marginal cost to the price level *without* rescaling by

$$\log P_t = \log \mathcal{P}_t \left(\{MC_s\} \right) = \log \mathcal{P}_t \left(\left\{ e^{\log MC_s} \right\} \right)$$

then the same mapping *with* rescaling is given by

$$\log P_t = \vartheta \log \mathcal{P}_t \left(\left\{ e^{\vartheta^{-1} \log MC_s} \right\} \right)$$

The partial derivative of $\log P_t$ to $\log MC_s$, evaluated in the model *with* rescaling around the steady state where $\log MC_{ss} = 0$, is then

$$\frac{\partial \log P_t}{\partial \log MC_s} = \vartheta \frac{\partial \log \mathcal{P}_t}{\partial \log MC_s} \times \vartheta^{-1} = \frac{\partial \log \mathcal{P}_t}{\partial \log MC_s}$$

The pass-through matrix Ψ is unchanged by ϑ. It therefore only depends on the ratio ξ/σ_ϵ^2.

C Appendix to Section 3

C.1 Envelope results for proof of proposition 1

Starting around the steady state, if there is a contemporaneous shock dc where $c \equiv \log MC$, then clearly $\frac{dV(x)}{dc} = -x$ from (A.18). Then, letting $V_n(x)$ denote the value function with n periods of anticipation of this shock, so that $dV_0(x) = -x d\log MC$, (A.19) gives the recursion from dV_{n-1} to dV_n. Further, this recursion preserves the property that $dV_n(0) = 0$ always, and hence simplifies to just

$$dV_n(x) = \beta(1-\lambda) \int_{-\bar{x}}^{\bar{x}} f(x'-x) dV_{n-1}(x') dx' \tag{A.21}$$

which is equivalent to our envelope result (27).

Observing that (A.21) is exactly the recursion that defines $E^n(x)$ given the same base case (times -1) of $E^0(x) = x$, but with an extra β added on each iteration, we conclude that

$$\frac{dV_n(x)}{dc} = -\beta^n E^n(x) \tag{A.22}$$

and also

$$\frac{dV_n'(x)}{dc} = -\beta^n (E^n)'(x)$$

Furthermore, the recursion (A.21) is the same as the recursion (A.11) obeyed by the steady-state $V'(x)$, except that the latter adds x on every iteration (rather than just the base case). Since we have shown that (A.11) is a contraction with the norm (A.17), it follows that the steady state obeys both

$$V'(x) = \sum_{n=0}^{\infty} \beta^n E^n(x) \equiv F(x)$$

and

$$V''(x) = \sum_{n=0}^{\infty} \beta^n (E^n)'(x) \equiv F'(x)$$

which both converge uniformly.

C.2 Proof that virtual survival is positive and decreasing

For the time-dependent models characterized in proposition 1 to be standard time-dependent models, their survival functions (which we call *virtual* survival functions) must be nonnegative and non-increasing. Here, we will show that given our assumptions, they are indeed *strictly* positive and decreasing.

First, we introduce an operator that will be useful both here and in the proof of proposition 2 in the next subsection. Define $T : L^2([-\bar{x}, \bar{x}]) \to L^2([-\bar{x}, \bar{x}])$ by

$$(Tg)(x) = (1-\lambda) \int_{-\bar{x}}^{\bar{x}} f(x'-x) g(x') dx' \tag{A.23}$$

given free adjustment rate λ and shock density f, and some policy bands $\bar{x}$. Recall the conditions we assumed in section 2.1 on the density f: differentiable, symmetric around 0, and single-peaked at 0, with $f'(x) < 0$ for $x > 0$ and vice versa (implying that $f(x) > 0$ everywhere).[A-6]

We note that T can be used to define expectation functions, since the recursion (2.1) is equivalent to $E^n = TE^{n-1}$. It can also be used to iterate on the density of price gaps, dropping readjusters: given an end-of-period density ϕ^{n-1} yesterday, the end-of-period density of non-readjusting firms is $\phi^n = T\phi^{n-1}$ today.

Properties of the operator T. We can obtain from (A.23) several properties of T. Of these, a) through c) will be needed for this proof, while d) and e) will be used to characterize eigenvalues in the next section.

[A-6] These assumptions are essential for our result, and it is possible to break the result when they are violated. To take a simple example that retains symmetry but breaks single-peakedness, suppose that $f(x) = \frac{1}{2}\delta(x - 2\bar{x}) + \frac{1}{2}\delta(x + 2\bar{x})$, where δ is a Dirac delta. Here, without adjustment, x either declines by $2\bar{x}$ or increases by $2\bar{x}$. Hence, starting at $\bar{x}$, there is a $\frac{1}{2}$ chance of declining to $-\bar{x}$ and not adjusting, and a $\frac{1}{2}$ chance of increasing to $3\bar{x}$ and adjusting back to 0. Iterating forward, we see that $E^n(\bar{x}) = (-\frac{1}{2})^n \bar{x}$, implying a virtual survival function that is both sometimes negative and non monotonic.

a) T is a Hilbert-Schmidt integral operator on $L^2([-\bar{x}, \bar{x}])$ with symmetric kernel $k(x, x') = f(x' - x)$, and therefore is self-adjoint. Since the kernel is continuous, Tg is continuous for any g.

b) T maps even g to even g, and odd g to odd g.

c) For any odd g where $g(x) \geq 0$ for all $x \geq 0$, and where $g(x) > 0$ for some positive-measure subset, $Tg(x) > 0$ for all $x > 0$. To see this, for any $x > 0$ exploit oddness to write $(Tg)(x) = \int_0^{\bar{x}} (f(x' - x) - f(-x' - x))\, g(x')dx' > 0$, where the inequality follows because single-peakedness implies $f(x' - x) - f(-x' - x) > 0$ for all $x, x' > 0$, and g is strictly positive on some subset of $[0, \bar{x}]$ of positive measure.

d) Eigenfunctions ψ of T with nonzero eigenvalues are continuously differentiable. If (ψ, μ) be any eigenfunction-eigenvalue pair with $\mu \neq 0$, note that $\psi = \mu^{-1} T\psi$, and so ψ must be C^0 by a). Applying $\psi = \mu^{-1} T\psi$ and (A.23) again shows that it is C^1.

e) Eigenvalues of T must have magnitude strictly less than 1. To see this, let (ψ, μ) be any eigenfunction-eigenvalue pair with $\mu \neq 0$. From property d), ψ is continuous and attains its maximum on $[-\bar{x}, \bar{x}]$ at some x^*. Then we observe that $|\lambda||\psi(x^*)| = |(T\psi)(x^*)| \leq |\psi(x^*)| \int_{-\bar{x}}^{\bar{x}} f(x' - x^*)dx' < |\psi(x^*)|$, and hence $|\lambda| < 1$.

Characterizing the expectation functions. We note that $E^0(x) = x$ is odd and satisfies $E^0(x) > 0$ for all $x > 0$. Hence, applying b) and c) above, it follows recursively that $T^n E^0 = E^n$ is odd and satisfies $E^n(x) > 0$ for all $x > 0$. It also follows from a) that E^n is continuous.

We observe that $E^1(x) < x$ for all $0 < x \leq \bar{x}$:

$$\begin{aligned}
E^1(x) &= \int_{-\bar{x}}^{\bar{x}} f(x' - x)x'dx' \\
&= \int_{x-(\bar{x}-x)}^{x+(\bar{x}-x)} f(x' - x)x'dx' + \int_{-\bar{x}}^{x-(\bar{x}-x)} f(x' - x)x'dx' \\
&= x \int_{x-(\bar{x}-x)}^{x+(\bar{x}-x)} f(x' - x)dx' + \int_{-\bar{x}}^{x-(\bar{x}-x)} f(x' - x)x'dx' \\
&\leq x \int_{-\bar{x}}^{\bar{x}} f(x' - x)dx' < x
\end{aligned}$$

Here, the equality in the third line follows from the symmetry of f, and the first inequality in the fourth line follows from $x' \leq x$ on the interval $[-\bar{x}, x - (\bar{x} - x)]$.

Now, for $n \geq 1$ define $\phi_x^n(x')$ to be the density of price gaps assuming that the price gap n periods ago was x, dropping all firms that have adjusted prices. This can be defined recursively using the base case $\phi_x^1(x') = f(x' - x)$ and $\phi_x^n = T\phi_x^{n-1}$. We observe that $(\phi_x^1 - \phi_{-x}^1)(x') = f(x' - x) - f(x' + x)$ is an odd function by symmetry of f, and that it also satisfies $(\phi_x^1 - \phi_{-x}^1)(x') > 0$ for all $x' > 0$ by single-peakedness of f. Hence, like above with E^n, it follows recursively from b) and

c) that for all n, $\phi_x^n - \phi_{-x}^n$ is odd and satisfies $(\phi_x^n - \phi_{-x}^n)(x') > 0$ for all x'. Symmetry also implies that $\phi_{-x}^n(x') = \phi_x^n(-x')$.

Next, we note that we can write any E^n either directly as the mean of x' given $\phi_x^n(x')$

$$E^n(x) = \int_{-\bar{x}}^{\bar{x}} \phi_x^n(x')x'dx' = \int_0^{\bar{x}} (\phi_x^n(x') - \phi_{-x}^n(x'))x'dx' \tag{A.24}$$

or as the mean of the one-period-ahead expectation $E^1(x')$ given $\phi_x^{n-1}(x')$

$$E^n(x) = \int_{-\bar{x}}^{\bar{x}} \phi_x^{n-1}(x')E^1(x')dx' = \int_0^{\bar{x}} (\phi_x^{n-1}(x') - \phi_{-x}^{n-1}(x'))E^1(x')dx' \tag{A.25}$$

We can then combine (A.24), (A.25), and our result that $E^1(x') < x'$ for $0 < x' < \bar{x}$ write for any $0 < x \leq \bar{x}$:

$$\begin{aligned} E^n(x) &= \int_0^{\bar{x}} (\phi_x^{n-1}(x') - \phi_{-x}^{n-1}(x'))E^1(x')dx' \\ &< \int_0^{\bar{x}} (\phi_x^{n-1}(x') - \phi_{-x}^{n-1}(x'))x'dx' \\ &= E^{n-1}(x) \end{aligned} \tag{A.26}$$

Since extensive margin virtual survival is proportional to $E^n(\bar{x})$, (A.26) combined with our earlier result that $E^n(x)$ was positive for $x > 0$ implies that extensive margin virtual survival is strictly positive and strictly declining, as desired.

Intensive margin virtual survival is $(E^n)'(0)$. Differentiating $E^n(x) = \int_{-\bar{x}}^{\bar{x}} f(x' - x)E^{n-1}(x')dx'$ around $x = 0$, we obtain

$$\begin{aligned} (E^n)'(0) &= -\int_{-\bar{x}}^{\bar{x}} f'(x')E^{n-1}(x')dx' \\ &= 2\int_0^{\bar{x}} (-f'(x'))E^{n-1}(x')dx' \end{aligned} \tag{A.27}$$

We observe that (A.27) integrates $E^{n-1}(x')$ with weights $-f'(x')$ that do not depend on n and are strictly positive on the interval $0 < x' \leq \bar{x}$. We know from above that for $0 < x' \leq \bar{x}$, both $E^{n-1}(x') > 0$ and $E^n(x') < E^{n-1}(x')$. It follows that $(E^n)'(0)$ is strictly positive for all n and strictly declining in n, as desired.

C.3 Elementary proof that weights sum to 1[A-7]

From (32), we see that the two coefficients on the two TD pass-through matrices are

$$2(1-\lambda)g(\overline{x})\overline{x} \cdot \sum_{s \geq 0} \frac{E^s(\overline{x})}{\overline{x}} \tag{A.28}$$

[A-7] We are grateful to an anonymous referee for the proof in this section.

and

$$\text{freq} \cdot \sum_{s \geq 0} E^{s\prime}(0) \tag{A.29}$$

Here, we provide an elementary proof that these two terms add to 1, rather than relying on the argument given in the main text. We begin by rewriting the extensive margin coefficient (A.28), followed by rewriting the intensive margin coefficient (A.29). At the end, we combine both expressions to show that their sum indeed equals 1.

Along the way, we use two auxiliary equations. The first is the condition (5) for $g(x)$ to be the steady state distribution of price gaps,

$$g(x) = \text{freq} \cdot f(x) + (1-\lambda) \int_{\underline{x}}^{\overline{x}} f(x-\tilde{x}) g(\tilde{x}) d\tilde{x} \tag{A.30}$$

from which it also follows that

$$1 = \text{freq} + (1-\lambda) \int_{\underline{x}}^{\overline{x}} g(x) dx \tag{A.31}$$

by integrating both sides in (A.30). It also follows that

$$g'(x) = \text{freq} \cdot f'(x) + (1-\lambda) \int_{\underline{x}}^{\overline{x}} f'(x-\tilde{x}) g(\tilde{x}) d\tilde{x} \tag{A.32}$$

from taking the derivative.

The second is the recursive formula (19) for $E^s(x)$, which for $s > 0$ specifies that

$$E^s(x) = x + (1-\lambda) \int f(\tilde{x}-x) E^{s-1}(\tilde{x}) d\tilde{x} \tag{A.33}$$

We can apply (A.33) to the sum $\sum_{s \geq 0} E^{s\prime}(x)$ to find

$$\sum_{s \geq 0} E^{s\prime}(x) = 1 - (1-\lambda) \int f'(\tilde{x}-x) \sum_{s \geq 0} E^s(\tilde{x}) d\tilde{x} \tag{A.34}$$

where the minus sign enters from differentiating $f'(\tilde{x}-x)$ with respect to x.

Extensive margin coefficient. In a first step, we apply the fundamental theorem of calculus to write (A.28) using an integral

$$2(1-\lambda) g(\overline{x}) \left(\sum_{s \geq 0} E^s(\overline{x}) \right) = (1-\lambda) \int_{\underline{x}}^{\overline{x}} \left(g(x) \sum_{s \geq 0} E^s(x) \right)' dx$$

The product rule allows us to break this integral into two separate terms,

$$2(1-\lambda)g(\overline{x})\left(\sum_{s\geq 0}E^s(\overline{x})\right) = \underbrace{(1-\lambda)\int_{\underline{x}}^{\overline{x}} g'(x)\sum_{s\geq 0}E^s(x)dx}_{I_1} + \underbrace{(1-\lambda)\int_{\underline{x}}^{\overline{x}} g(x)\sum_{s\geq 0}E^{s\prime}(x)dx}_{I_2}$$

Each of these integrals can be rewritten further. We give them the names I_1 and I_2.

- Integral I_1 can be rewritten using (A.32) as

$$I_1 = (1-\lambda)\int_{\underline{x}}^{\overline{x}} \text{freq}\cdot f'(x)\sum_{s\geq 0}E^s(x)dx + (1-\lambda)^2\int_{\underline{x}}^{\overline{x}}\int_{\underline{x}}^{\overline{x}} f'(x-\tilde{x})g(\tilde{x})d\tilde{x}\sum_{s\geq 0}E^s(x)dx \tag{A.35}$$

- Integral I_2 can be rewritten using (A.34) as

$$I_2 = (1-\lambda)\int_{\underline{x}}^{\overline{x}} g(x)dx - (1-\lambda)^2\int_{\underline{x}}^{\overline{x}} g(x)\int_{\underline{x}}^{\overline{x}} f'(\tilde{x}-x)\sum_{s\geq 0}E^s(\tilde{x})d\tilde{x}dx \tag{A.36}$$

Adding I_1 and I_2 we find that the double integrals in (A.35) and (A.36) exactly cancel by simply relabeling x and $\tilde{x}$. We thus arrive at the following expression for the extensive-margin coefficient

$$2(1-\lambda)g(\overline{x})\left(\sum_{s\geq 0}E^s(\overline{x})\right) = I_1 + I_2 = (1-\lambda)\int_{\underline{x}}^{\overline{x}} \text{freq}\cdot f'(x)\sum_{s\geq 0}E^s(x)dx + (1-\lambda)\int_{\underline{x}}^{\overline{x}} g(x)dx$$

Intensive margin coefficient. We rewrite the intensive margin coefficient (A.29) using (A.34)

$$\text{freq}\cdot\sum_{s\geq 0}E^{s\prime}(0) = \text{freq} - \text{freq}\,(1-\lambda)\int_{\underline{x}}^{\overline{x}} f'(x)\sum_{s\geq 0}E^s(x)dx$$

Summing the coefficients. Summing the two coefficients, we find that the two remaining terms that involve $E^s(x)$ cancel,

$$2(1-\lambda)g(\overline{x})\left(\sum_{s\geq 0}E^s(\overline{x})\right) + \text{freq}\cdot\sum_{s\geq 0}E^{s\prime}(0) = \text{freq} + (1-\lambda)\int_{\underline{x}}^{\overline{x}} g(x)dx$$

which, given (A.31), is exactly equal to 1.

C.4 Proof of proposition 2

Applying the spectral theorem. Since T is self-adjoint and (like all Hilbert-Schmidt integral operators) compact from property a), we can apply the spectral theorem for separable infinite-dimensional Hilbert spaces, which states that $L^2([-\bar{x},\bar{x}])$ has a countably infinite orthonormal

basis $\{\psi_n\}$ of eigenfunctions of T, with corresponding real eigenvalues $\{\mu_n\}$, where the only accumulation point of μ_n is 0.

Indeed, since by property b), T preserves evenness and oddness, we can also define it on the even and odd subspaces of $L^2([-\bar{x}, \bar{x}])$, and then apply the spectral theorem separately on each subspace, to get separate orthonormal bases for the even subspace and the odd subspace. Since the sum of the even and odd subspaces is the entire function space, these bases combine to form an orthonormal basis for all of $L^2([-\bar{x}, \bar{x}])$. Hence, we can further refine our statement in the first paragraph, and conclude that all eigenfunctions ψ_n in the orthonormal basis are either even or odd.

Extensive and intensive margin survival curves. Using the above, we project E^0 onto the basis of eigenfunctions

$$E^0 = \sum_n \langle E^0, \psi_n \rangle \psi_n \tag{A.37}$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on L^2. Note that $\langle E^0, \psi_n \rangle$ will be zero for all even ψ_n, and only nonzero for some odd ψ_n.

Applying T to (A.37) gives

$$E^s(x) = \sum_n \langle E^0, \psi_n \rangle \mu_n^s \psi_n(x) \tag{A.38}$$

which now holds point-wise for any $s > 0$.[A-8]

Define $\bar{\mu}$ to be the eigenvalue μ_n of maximum magnitude such that $\langle E^0, \psi_n \rangle$ is nonzero, and define $\bar{\psi}$ to be the projection of E^0 onto the corresponding eigenspace, i.e.

$$\bar{\psi}(x) \equiv \sum_{\{n:\mu_n = \bar{\mu}\}} \langle E^0, \psi_n \rangle \psi_n(x)$$

where we note that $\bar{\psi}(x)$ is odd. Then it follows from (A.38) that

$$\lim_{s \to \infty} \frac{E^s(x)}{\bar{\mu}^s} = \bar{\psi}(x) \tag{A.39}$$

Since $E^0(x) = x > 0$ for all $x > 0$, repeatedly applying property c) it must also be true that $E^s(x) > 0$ for all $x > 0$. Taking the limit (A.39), we must have $\bar{\psi}(x) \geq 0$ for all $x > 0$ and also $\bar{\mu} > 0$. Further, given that $\bar{\psi}$ is continuous and by construction is not identically zero, we must have $\bar{\psi}(x) > 0$ for some subset of $[0, \bar{x}]$ of positive measure, and so $(T\bar{\psi})(x) = \lambda \bar{\psi}(x) > 0$ for all $x > 0$. We conclude that $\bar{\psi}(x) > 0$ for all $x > 0$.

[A-8] Normally this projection could differ on a set of measure 0, but from properties 1 and 3 we know that $E^1 = TE^0$ is continuous and also that ψ_n is continuous for any $\mu_n \neq 0$, and continuous functions cannot differ on a set of only measure 0.

Using this result, we can also write

$$\begin{aligned}\mu\bar{\psi}'(0) = (T\bar{\psi})'(0) &= -\int_{-\bar{x}}^{\bar{x}} \phi'(x')\bar{\psi}(x')dx' \\ &= -2\int_{0}^{\bar{x}} \phi'(x')\bar{\psi}(x')dx' > 0\end{aligned}$$

and we conclude that $\bar{\psi}'(0) > 0$ as well.

The intensive and extensive margin "virtual survival" curves are given by $\Phi_t^i \equiv (E^t)'(0)$ and $\Phi_t^e \equiv E^t(\bar{x})/\bar{x}$, which using (A.39) and the preceding results have the limits

$$\begin{aligned}\lim_{s\to\infty} \frac{\Phi_t^i}{\bar{\mu}^s} &= \bar{\psi}'(0) > 0 \\ \lim_{s\to\infty} \frac{\Phi_t^e}{\bar{\mu}^s} &= \frac{\bar{\psi}(\bar{x})}{\bar{x}} > 0\end{aligned}$$

It follows that both Φ_t^i and Φ_t^e asymptotically decay at the rate $\bar{\mu}$, and that their asymptotic hazards are both $1-\bar{\mu}$.

The asymptotic hazard of virtual survival is strictly greater than that of actual survival. The subset of weakly positive functions is a total cone in $L^2([-\bar{x},\bar{x}])$, and from property c), applying T to any of these functions that is nonzero gives a function that is strictly positive everywhere and therefore in the interior of the cone. In other words, T is strongly positive with respect to this cone, and we can apply a standard extension of the Krein-Rutman theorem[A-9] to conclude that the eigenvalue $\hat{\mu}$ of T with largest magnitude is simple, strictly positive, and has a corresponding eigenfunction $\hat{\psi}$ that is strictly positive. Since $\hat{\psi}$ is strictly positive, we note that it must be even rather than odd.

Now, let the probability of survival s periods from now starting at some point x be $\Phi_s^{actual}(x)$, where $\Phi_0^{actual}(x) \equiv 1$ and $\Phi_s^{actual} = T^s\Phi_0^{actual}$. We know that

$$\Phi_s^{actual}(x) = \sum_n \langle \Phi_0^{actual}, \psi_n\rangle \mu_n^s \psi_n(x)$$

and note that since $\hat{\psi}$ is strictly positive, $\langle \Phi_0^{actual}, \hat{\psi}\rangle > 0$, so that taking the limit as $s \to \infty$ we have

$$\lim_{s\to\infty} \frac{\Phi_s^{actual}(x)}{\hat{\mu}^s} = \langle \Phi_0^{actual}, \hat{\psi}\rangle \hat{\psi}(x)$$

so that the asymptotic hazard rate of actual survival is $1-\hat{\mu}$.

Since $\hat{\mu}$ is the (simple) dominant eigenvalue and corresponds to an even eigenfunction, it must be strictly larger than $\bar{\mu}$, which was associated with odd eigenfunctions. Hence the asymptotic hazard of actual survival, $1-\hat{\mu}$, is less than that of virtual survival, $1-\bar{\mu}$.

[A-9] See Theorem 19.3 in Deimling (1985).

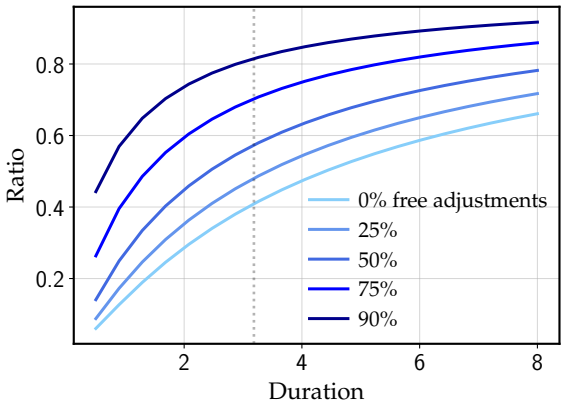


Figure C.1: Ratio of second to first odd eigenvalues of the transition operator across the parameter space of the menu cost model

C.5 State-dependent models that are exactly equivalent to Calvo

Here, we revisit the question of whether there are more SD models that are exactly equivalent to Calvo models, in the spirit of Gertler and Leahy (2008).

To investigate this, we look for densities $f(\epsilon)$ of the idiosyncratic shock ϵ_{it} that generate $E^t(x) = \phi^t x$ for some $\phi \in [0,1)$ and for all $x \in [\underline{x}, \overline{x}]$. This is sufficient, but not necessary for Calvo, since Calvo only requires that $E^{t\prime}(0) = \phi^t = E^t(\overline{x})/\overline{x}$. Moreover, notice $E^t(x) = \phi^t x$ follows from $E^1(x) = \phi x$ by induction since if $E^t(x) = \phi^t x$ holds, then it is also the case that

$$E^{t+1}(x) = \int_{\underline{x}}^{\overline{x}} f(x'-x)E^t(x')dx' = \int_{\underline{x}}^{\overline{x}} f(x'-x)\phi^t x' dx' = \phi^t E^1(x) = \phi^{t+1}x$$

So which densities $f(\epsilon)$ guarantee that $E^1(x) = \phi x$? It needs to be the case that

$$\phi x = \int_{\underline{x}}^{\overline{x}} f(x'-x)x'dx'$$

for any $x \in [\underline{x}, \overline{x}]$. Taking derivatives and integrating by parts, this implies

$$\phi + f(\overline{x}-x)\overline{x} - f(\underline{x}-x)\underline{x} = F(\overline{x}-x) - F(\underline{x}-x) \tag{A.40}$$

where we denote the cdf of f by F. Taking derivatives one more time, we find

$$-f'(\overline{x}-x)\overline{x} + f'(\underline{x}-x)\underline{x} = -f(\overline{x}-x) + f(\underline{x}-x)$$

Using $\underline{x} = -\overline{x}$ and the fact that f is symmetric and f' is anti-symmetric, we find

$$f(\overline{x}-x) - f'(\overline{x}-x)\overline{x} = f(\overline{x}+x) - f'(\overline{x}+x)\overline{x} \tag{A.41}$$

This equation has to hold for any $x \in [0, \overline{x}]$.

Observe that without loss, we can normalize $\overline{x}$ to 1. Why? Because if and only if we find a density $\hat{f}(\epsilon)$ for which (A.41) holds with $\overline{x} = 1$, then $f(\epsilon) \equiv \hat{f}(\epsilon/\overline{x})$ satisfies (A.41) with any other $\overline{x} > 0$. With $\overline{x} = 1$, (A.41) is

$$f(1-x) - f'(1-x) = f(1+x) - f'(1+x) \tag{A.42}$$

Now, let us define the following function: $g(\epsilon) \equiv e^{-\epsilon} f(\epsilon)$ for $\epsilon \in [0,2]$. Its derivative is equal to

$$g'(\epsilon) = -e^{-\epsilon}\left(f(\epsilon) - f'(\epsilon)\right)$$

which is convenient since terms involving $f - f'$ are exactly what appears in (A.42). In particular, we can write

$$\begin{aligned} g'(1+x) &= -e^{-(1+x)}\left(f(1+x) - f'(1+x)\right) \\ g'(1-x) &= -e^{-(1-x)}\left(f(1-x) - f'(1-x)\right) \end{aligned}$$

so that (A.42) can be rewritten as

$$g'(1-x) = e^{2x} g'(1+x) \tag{A.43}$$

Any positive differentiable g, defined on $[0,2]$ that satisfies (A.43) gives us a density $f(\epsilon) = e^{\epsilon} g(\epsilon)$ (up to scale) of an SD model that is exactly equivalent to Calvo.

A simple example. We guess $g'(1+x) = -be^{-c(1+x)}$ for some constants $c \in (0,1]$ and $b > 0$. Then, by (A.43) it has to be that

$$g'(1-x) = -be^{-c+(2-c)x}$$

Integrating g', we therefore find

$$g(\epsilon) = \begin{cases} a + \frac{b}{2-c} e^{2-2c-(2-c)\epsilon} & \epsilon \leq 1 \\ a + \frac{b}{2-c} e^{-c} - \frac{b}{c}\left(e^{-c} - e^{-c\epsilon}\right) & \epsilon > 1 \end{cases}$$

where for simplicity of this example we set the constant $a = 0$. Multiplying with e^{ϵ} gives the density f for $\epsilon \geq 0$ (and symmetrically for $\epsilon \leq 0$)

$$f(\epsilon) = \begin{cases} \frac{b}{2-c} e^{2-2c-(1-c)\epsilon} & \epsilon \leq 1 \\ \frac{b}{2-c} e^{-c+\epsilon} - \frac{b}{c}\left(e^{-c} - e^{-c\epsilon}\right) e^{\epsilon} & \epsilon > 1 \end{cases}$$

Using this expression, we see that f has positive support on $[-\overline{y}, \overline{y}]$, where $\overline{y} = 1 - \frac{1}{c}\log\frac{2-2c}{2-c}$. To normalize f, we can choose b such that $2\int_0^{\overline{y}} f(y)dy = 1$. This gives a closed form expression

for b. And with that, we can compute the Calvo hazard $1-\phi$, where ϕ follows from (A.40)

$$\phi = 2\frac{b}{2-c}\frac{1}{1-c}e^{2-2c}\left[1-(2-c)\,e^{-(1-c)}\right]$$

where we simply used (A.40) with $x=0$ (any value of x works).

C.6 Exact equivalence between SD and TD models with trend inflation

In this appendix, we state and prove a generalization of proposition 1 for the case in which there is trend inflation, i.e., we add a drift term μ to the law of motion for the static optimal price (1), which becomes:

$$p^*_{it} = p^*_{it-1} + \mu + \epsilon_{it}.$$

We have the following result:

Proposition 7. *The pass-through matrix Ψ of the canonical menu cost model with trend inflation is a mixture of three time-dependent pass-through matrices,*

$$\Psi = \alpha^{\ell}\Psi^{\Phi^{\ell}} + \alpha^{u}\Psi^{\Phi^{u}} + (1-\alpha^{\ell}-\alpha^{u})\Psi^{\Phi^{i}},$$

associated with the survival functions

$$\Phi^{\ell}_s = \frac{E^s(\underline{x}) - E^s(x^*)}{\underline{x} - x^*}$$

$$\Phi^{u}_s = \frac{E^s(\overline{x}) - E^s(x^*)}{\overline{x} - x^*}$$

$$\Phi^{i}_s = (E^s)'(x^*).$$

Now, we need three time-dependent components because the model is no longer symmetric. Without trend inflation, the lower and upper adjustment boundaries move symmetrically, so their dynamics coincide and can be captured by a single TD model. With $\mu \neq 0$, we need one TD model for the lower adjustment boundary $\Psi^{\Phi^{\ell}}$ and one for the upper boundary $\Psi^{\Phi^{u}}$, as well as one for the reset point, or intensive margin, $\Psi^{\Phi^{i}}$.

The proof of this proposition is very similar to the one of proposition 1, but with minor adjustments to take care of the fact that the steady state is no longer symmetric, i.e. $x^* \neq 0$. We follow the same steps as in section 3.2.

Expected price gaps. Here, we still define $E^t(x) = \mathbb{E}_0\left[x_{it}|x_{i0}=x\right]$. However, the recursion that characterizes it needs to be slightly adjusted. We have

$$E^t(x) = (1-\lambda)\int_{\underline{x}}^{\overline{x}} f(x'-x)E^{t-1}(x')dx' + \text{freq}\cdot E^{t-1}(x^*), \tag{A.44}$$

where freq is the steady state adjustment frequency.

Relationship between expected price gaps and the law of motion. With trend inflation, equation (23) still holds, but totally differentiating it around steady state gives us

$$\begin{aligned} d\log P_t &= (1-\lambda)g(\underline{x})\sum_{s=0}^{\infty}\left(E^s(\underline{x})-E^s(x^*)\right)d\underline{x}_{t-s} \\ &+ (1-\lambda)g(\overline{x})\sum_{s=0}^{\infty}\left(E^s(\overline{x})-E^s(x^*)\right)d\overline{x}_{t-s} \\ &+ \text{freq}\sum_{s=0}^{\infty}\left(E^s\right)'(x^*)dx^*_{t-s}, \end{aligned} \tag{A.45}$$

which is analogous to (25). The evolution of the price level after a change in adjustment boundaries now depends on the difference between the expected price gaps at the boundaries and at the reset point, $E^s(\underline{x})-E^s(x^*)$ and $E^s(\overline{x})-E^s(x^*)$. In the symmetric case, $x^*=0$ and $E^s(x^*)=0$, which do not necessarily hold when $\mu\neq 0$.

Relationship between expected price gaps and the policy equation. The recursive formulation (26) for the value function $V_t(x)$ still holds with trend inflation. From this, we obtain (27) and (28) using the same steps as before. Moreover, it is also the case that $V'(x)=\sum_{t=0}^{\infty}\beta^t E^t(x)$, which for $\mu=0$ we prove in appendix C.1. With trend inflation, however, this requires a slightly modified proof, which we provide here.

As in appendix C.1, define

$$F(x)=\sum_{t=0}^{\infty}\beta^t E^t(x).$$

$F(x)$ is characterized by the following recursion:

$$F(x)=x+\beta\left[(1-\lambda)\int_{\underline{x}}^{\overline{x}}f(x'-x)F(x')dx'+\left(\lambda+(1-\lambda)\left(1-\int_{\underline{x}}^{\overline{x}}f(x'-x)dx'\right)\right)F(x^*)\right]. \tag{A.46}$$

By following the same steps as in appendix C.1, we obtain the same recursion for $V'(x)$ as in the zero inflation case:

$$V'(x)=x+\beta(1-\lambda)\int_{\underline{x}}^{\overline{x}}V'(x')f(x'-x)dx'. \tag{A.47}$$

In order to prove that $V'(x)=F(x)$, we show that $F(x^*)=0$, from which it follows that $V'(x)$ and $F(x)$ fixed points of the same contraction operator.

To see why $F(x^*) = 0$, subtract $F(x^*)$ from (A.46) to obtain

$$F(x) - F(x^*) = x + \beta(1-\lambda)\int_{\underline{x}}^{\overline{x}} f(x'-x)\left(F(x') - F(x^*)\right)dx' - (1-\beta)F(x^*) \tag{A.48}$$

Now iterate both (A.47) and (A.48) forward to obtain

$$V'(x) = F(x) - F(x^*) + (1-\beta)F(x^*)\sum_{t=0}^{\infty}\beta^t p_t(x),$$

where $p_t(x)$ is the probability of not having adjusted t periods into the future conditional on starting at x. Plug $x = x^*$ into this equation and use $V'(x^*) = 0$ to obtain $F(x^*) = 0$.

Finally, one can proceed exactly as in section 3 to obtain

$$d\overline{x}_t = -\frac{(dV_t(\overline{x}) - dV_t(x^*))}{V'(\overline{x})}$$

$$d\underline{x}_t = -\frac{(dV_t(\underline{x}) - dV_t(x^*))}{V'(\overline{x})},$$

from which it follows that

$$d\overline{x}_t = \frac{\sum_{s\geq t}\beta^{s-t}\left(E^{s-t}(\overline{x}) - E^{s-t}(x^*)\right)d\log MC_s}{\sum_{s\geq t}\beta^{s-t}\left(E^{s-t}(\overline{x}) - E^{s-t}(x^*)\right)}$$

$$d\underline{x}_t = \frac{\sum_{s\geq t}\beta^{s-t}\left(E^{s-t}(\underline{x}) - E^{s-t}(x^*)\right)d\log MC_s}{\sum_{s\geq t}\beta^{s-t}\left(E^{s-t}(\underline{x}) - E^{s-t}(x^*)\right)}.$$

For the reset point, we obtain

$$dx_t^* = \frac{\sum_{s\geq t}\beta^{s-t}\left(E^{s-t}\right)'(x^*)d\log MC_s}{\sum_{s\geq t}\beta^{s-t}\left(E^{s-t}\right)'(x^*)}.$$

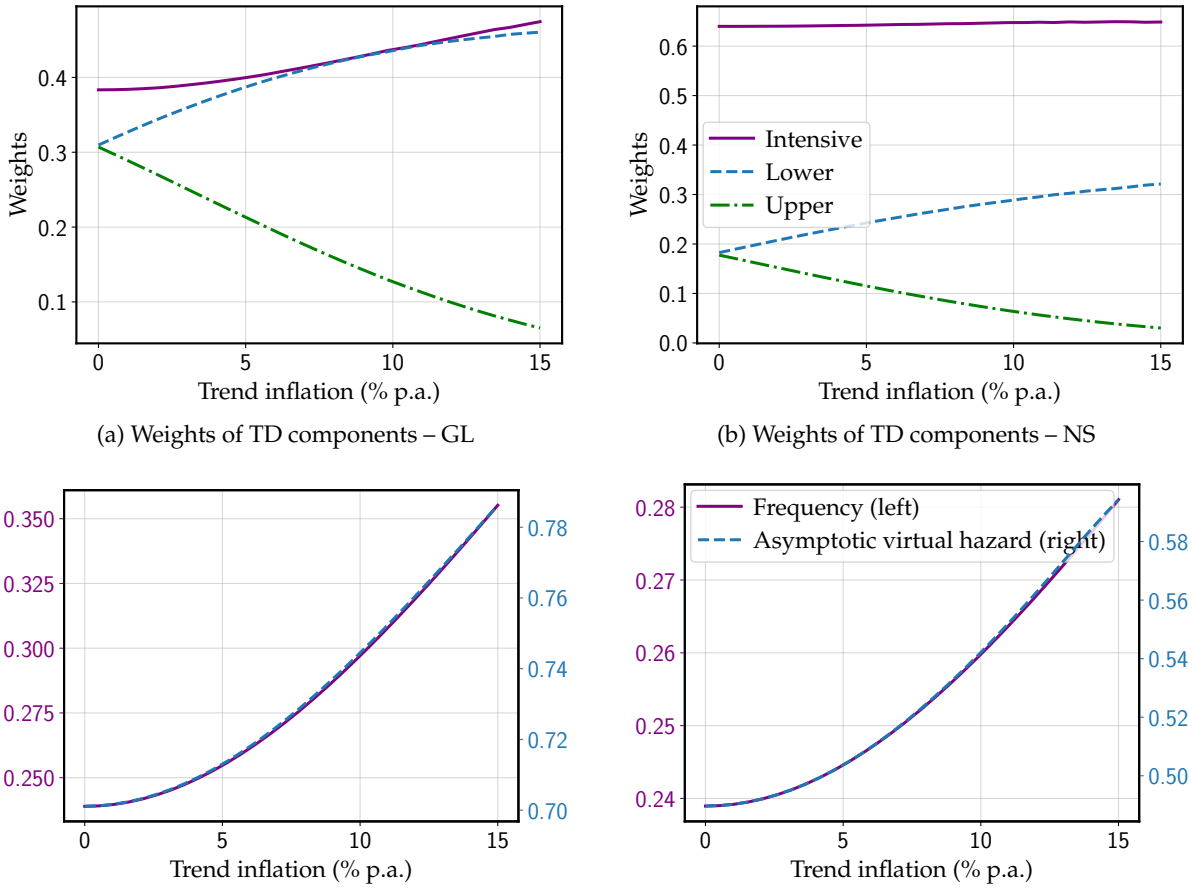
Proposition 7, therefore, follows as in section 3 by defining

$$\alpha^{\ell} = (1-\lambda)g(\underline{x})\sum_{s=0}^{\infty}\left(E^s(\underline{x}) - E^s(x^*)\right)$$

$$\alpha^{u} = (1-\lambda)g(\overline{x})\sum_{s=0}^{\infty}\left(E^s(\overline{x}) - E^s(x^*)\right)$$

$$\alpha^{i} = 1 - \alpha^{\ell} - \alpha^{u} = \text{freq}\sum_{s=0}^{\infty}\left(E^s\right)'(x^*).$$

The top two graphs of figure C.2, panels (a) and (b), show the weight of each of the three time-dependent components as we increase trend inflation, for the GL and NS calibrations. The results are intuitive: as inflation increases, the weight of the lower bound increases, as there are more up-



(a) Weights of TD components – GL

(b) Weights of TD components – NS

(c) Frequency and asymptotic virtual hazard – GL

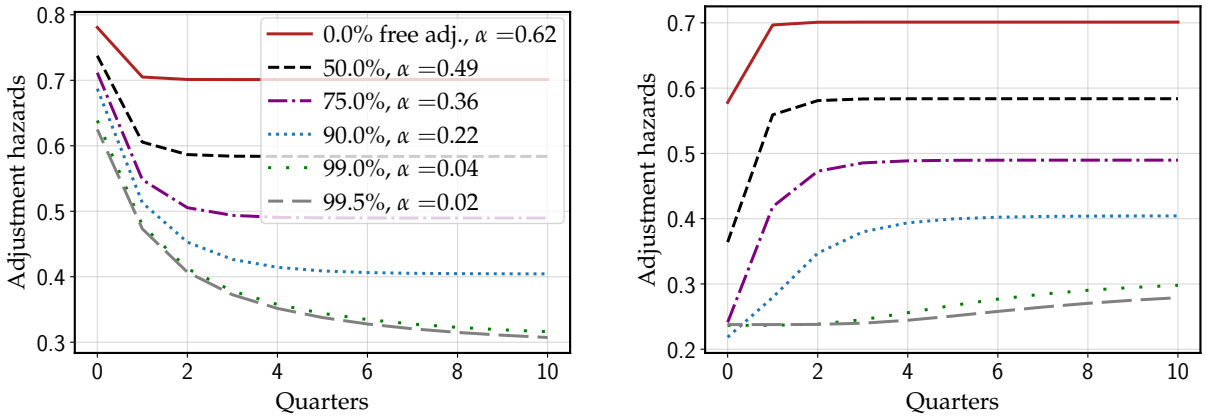
(d) Frequency and asymptotic virtual hazard – NS

Figure C.2: Weights of each TD component, adjustment frequencies, and asymptotic adjustment hazards for various levels of trend inflation

Note: the top two panels show the weights of the TD equivalent components as a function of trend inflation, α^i, α^ℓ, and α^u, as characterized in the text. The bottom two panels show adjust frequencies for GL and NS models calibrated as in section 2.5 for various levels of trend inflation, as well as the common asymptotic adjustment hazards of the associated TD components.

ward adjustments, while the opposite happens for the weight of the upper boundary. The bottom two graphs, panels (c) and (d), show adjustment frequencies and asymptotic virtual hazards, i.e., the common limiting value of the virtual adjustment hazards of the time-dependent components. Again intuitively, the adjustment frequency increases with trend inflation, as the drift in price gaps pushes firms outside their inaction regions. Moreover, the asymptotic adjustment hazards increase, reflecting the resulting extra price flexibility that is generated by trend inflation.

Interestingly, the intensive margin weight has a slope of zero at zero trend inflation. Likewise, the slopes of the two extensive margin weights add to zero at zero trend inflation. This is part of a more general symmetry (see also Alvarez et al. 2016 among others). Using the above proof techniques, one can show that the entire pass-through matrix Ψ is, in fact, symmetric in trend inflation, so trend inflation only has a second order effect on inflation dynamics.



(a) Extensive margin hazards

(b) Intensive margin hazards

Figure C.3: Extensive and intensive margin hazards with varying shares of free adjustments

Note. These panels compute extensive and intensive margin hazards, as in figure 6, varying the share of free adjustments but holding the frequency constant. 0% corresponds to Golosov-Lucas; 75% to Nakamura-Steinsson.

C.7 Extensive and intensive margin hazards with high shares of free adjustments

We found in section 4.2 that the aggregate predictions of the GL model are more similar to Calvo than the aggregate predictions of the NS model, despite the greater share of free adjustments of that model. Figure C.3 illustrates that there are two opposing forces that drive this.

On the one hand, with a greater share of free adjustments, the weight α on the extensive margin is lower. Together with the intensive margin hazards flattening out, this rationalizes why, eventually, as the share of free adjustments goes to one, the model approaches a Calvo model.

On the other hand, a greater share of free adjustments also leads extensive margin hazards to become steeper. This is because, holding the frequency constant, a greater share of free adjustments is achieved by a wider *Ss* band; and a wider *Ss* band makes it less likely that price gaps at one of the adjustment boundaries mix with other price gaps, increasing the selection effect.

For moderate shares of free adjustments, the second effect dominates, making models with greater shares of free adjustments counterintuitively *less* Calvo-like. Eventually, as the share of free adjustments approaches one, the first effect takes over, and the model converges to a Calvo model with adjustment hazard λ, the free adjustment hazard.

D Appendix to Section 4

D.1 NKPC regressions with simulated data

In this section, we use simulated data from our two baseline SD models – Golosov-Lucas and Nakamura-Steinsson – to run the standard NKPC regression below:

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \gamma \pi_{t-1} + \kappa \hat{mc}_t + u_t \tag{A.49}$$

	Golosov-Lucas				Nakamura-Steinsson			
	K approx.	$\beta = 0.99$ $\gamma = 0$	$\gamma = 0$	Unrestricted	**K** approx.	$\beta = 0.99$ $\gamma = 0$	$\gamma = 0$	Unrestricted
κ	1.709	1.743	1.762	1.763	0.468	0.454	0.472	0.472
β	0.99	0.99	0.979	0.981	0.99	0.99	0.949	0.935
γ	0	0	0	-0.002	0	0	0	0.014
R^2	1.000	1.000	1.000	1.000	0.998	0.998	0.999	0.999

Table D.1: Regression results with simulated data.

The simulation procedure works as follows. We first posit a stochastic process for the real marginal cost $\hat{mc}_t$. Then, using the generalized Phillips curve (16), it is straightforward to jointly simulate paths for $\hat{mc}_t$, inflation π_t, lagged inflation π_{t-1}, and inflation expectations $\mathbb{E}_t\pi_{t+1}$. We then use these simulated paths to estimate equation (A.49) via ordinary least squares. We simulate a sample of size 10,000.

We assume that $\hat{mc}_t$ is given by the sum of an AR(1) process with persistence 0.8 and an i.i.d. shock term, both with the same unconditional variance.[A-10] For each SD model, we analyze four different specifications of (A.49). First, we fix β and κ at the values suggested by approximating the generalized Phillips curve **K**, as in the main text, and compute only the R^2 of the fit. Then, we fix $\beta = 0.99$ and $\gamma = 0$, and estimate only the slope κ. In the third specification, we fix $\gamma = 0$ and estimate β and κ. Finally, in the last specification we impose no parameter restrictions.

Table D.1 shows results. Standard errors of parameter estimates are negligible, and therefore omitted. There are several interesting features in these results. First, the estimated values of κ are close to the ones suggested by approximating the whole generalized Phillips curve, which is not surprising. Second, the estimated forward coefficient β may differ from the value 0.99 used in simulations, and the backward coefficient may slightly differ from zero when unrestricted. Finally, all regressions have very high R^2. It is not surprising, however, that the **K** approximation does not generate the highest R^2. This approximation maximizes the minimum R^2 of an NKPC regression among all possible finite-variance $\hat{mc}_t$ stochastic processes, which is not necessarily attained for a process similar to the one we use in this simulation exercise. Nevertheless, these simulations provide a concrete example of the extremely good fit of the Calvo approximation for SD models.

D.2 Approximate equivalence to Calvo when the average survival function is exponential

Define $\mathcal{L}(\{a^s\})$, for any sequence $\{a^s\}$, to be the lower triangular Toeplitz matrix with first column $\{a^s\}$, and similarly define $\mathcal{U}(\{a^s\})$ to be the upper triangular Toeplitz matrix with first row $\{a^s\}$. In this notation the definition (12) of pass-through matrix for a TD model, for instance, is $\left(\sum_s \Phi_s\right)^{-1}\left(\sum_s \beta^s\Phi_s\right)^{-1}\mathcal{L}(\{\Phi_s\})\mathcal{U}(\{\beta^s\Phi_s\})$.

[A-10] In a Calvo model, having only an AR(1) component generates perfect multicollinearity between $\mathbb{E}_t\pi_t$ and $\hat{mc}_t$. This problem extends to our SD models, hence the need for the i.i.d. term.

The equivalence result (17) can be written in this notation as

$$\Psi = \alpha \frac{\mathcal{L}(\{\Phi_s^e\})\mathcal{U}(\{\Phi_s^e\})}{(\sum_s \Phi_s^e)\,(\sum_s \beta^s \Phi_s^e)} + (1-\alpha)\frac{\mathcal{L}(\{\Phi_s^i\})\mathcal{U}(\{\Phi_s^i\})}{(\sum_s \Phi_s^i)\,(\sum_s \beta^s \Phi_s^i)} \tag{A.50}$$

Now, suppose that $\Phi_s^e = \Phi_s^{\text{Calvo}} + \eta_s^e$ and $\Phi_s^i = \Phi_s^{\text{Calvo}} + \eta_s^i$, with η_s^e and η_s^i small and $\Phi_s^{\text{Calvo}} \equiv (1-\lambda)^s$, i.e. that both the extensive and intensive margin virtual survival functions are close to exponential (Calvo). Then to first order in the ηs, (A.50) can be approximated as

$$\begin{aligned}\Psi \approx \left(1 - \frac{\sum_s \alpha\eta_s^e + (1-\alpha)\eta_s^i}{\sum_s \Phi_s^{\text{Calvo}}} - \frac{\sum_s \beta^s(\alpha\eta_s^e + (1-\alpha)\eta_s^i)}{\sum_s \beta^s \Phi_s^{\text{Calvo}}}\right)\Psi^{\text{Calvo}} \\ + \frac{\mathcal{L}(\{\alpha\eta_s^e + (1-\alpha)\eta_s^i\})\mathcal{U}(\{\Phi_s^{\text{Calvo}}\}) + \mathcal{L}(\{\Phi_s^{\text{Calvo}}\})\mathcal{U}(\{\alpha\eta_s^e + (1-\alpha)\eta_s^i\})}{(\sum_s \Phi_s^{\text{Calvo}})\,(\sum_s \beta^s \Phi_s^{\text{Calvo}})}\end{aligned} \tag{A.51}$$

where Ψ^{Calvo} is the Calvo pass-through matrix associated with Φ_s^{Calvo}.

Note that in (A.51), the η_s^e and η_s^i only appear together in the form $\alpha\eta_s^e + (1-\alpha)\eta_s^i$, which equals the discrepancy between the *average* survival function of the mixture and the Calvo survival function:

$$\alpha\eta_s^e + (1-\alpha)\eta_s^i = \alpha\Phi_s^e + (1-\alpha)\Phi_s^i - \Phi_s^{\text{Calvo}} \tag{A.52}$$

If the average survival function equals the Calvo exponential survival function, then to first order in the ηs the pass-through matrix of the menu cost model is the *same* as Calvo.

More generally, to first order in the ηs, the discrepancy between the menu cost pass-through matrix and the Calvo pass-through matrix is of the same magnitude as the gap between the average survival function of the mixture and the Calvo exponential survival function. If this gap is small (and the ηs are not too big), then the two pass-through matrices should be close, which carries over to the generalized Phillips curve.

D.3 Proof of the convergence of $\sum_k \Psi^k$

We offer two proofs of the convergence of $\sum_k \Psi^k$ in this paper. The first proof, which we lay out in this section, relies on two simple properties of Ψ and follows directly from established mathematical results. The second proof follows as a by-product of the proof of proposition 3 and is explained in appendix D.4.

Our first proof builds on the following very useful result. Denote by $\mathbb{N}_0 \equiv \{0, 1, \cdots\}$ the set of all natural numbers including 0.

Proposition 8. *Let Ψ be a positive $\mathbb{N}_0 \times \mathbb{N}_0$ matrix whose columns sum to weakly less than 1, and at least one column sums to strictly less than 1. Then, there exists a positive $\mathbb{N}_0 \times \mathbb{N}_0$ matrix $\mathbf{N}$ such that:*

a) $\mathbf{N}$ is the (element-wise) limit of $\Psi + \Psi^2 + \Psi^3 + \ldots$

b) all entries of $\mathbf{N}$ are finite.

c) $\mathbf{N} = \Psi + \Psi\mathbf{N}$, $\mathbf{N} = \Psi + \mathbf{N}\Psi$ *and thus* $(\mathbf{I} - \Psi)\,\mathbf{N} = \mathbf{N}\,(\mathbf{I} - \Psi) = \Psi$.

in other words, $\mathbf{N} = \sum_{k=1}^{\infty} \Psi^k$ *is the inverse of* $\mathbf{I} - \Psi$.

Proof. Let $\mathbf{P}$ denote the transpose of a matrix Ψ with those properties. $\mathbf{P}$ is a positive row sub-stochastic matrix with at least one row sum strictly below 1. Such a matrix can be viewed as the transition matrix of a transient Markov chain. This allows us to apply the results of chapter 5 in Kemeny, Snell and Knapp (1976). Their proposition 5-2 establishes that $\mathbf{N}$ is the element-wise limit of $\Psi + \Psi^2 + \ldots$. Proposition 5-3 establishes that $\mathbf{N}$ is finite valued. Proposition 5-4 establishes the properties listed under (c). □

It is left to show that the pass-through matrices Ψ in our applications satisfy the assumptions required for the application of proposition 8. But this is straightforward:

- For a TD model, Ψ is given by (12) and thus clearly positive for any positive survival rate schedule $\Phi_t > 0$ and positive discount factor $\beta > 0$.[A-11]
- For the pass-through matrix Ψ of our canonical menu cost model, we simply apply the result in proposition 1 to express Ψ as the convex combination of two TD pass-through matrices. This immediately establishes that Ψ satisfies the conditions of proposition 8.

D.4 Proof of proposition 3

In this appendix section, we prove proposition 3. Along the way, we will provide a second way to prove that the infinite sum $\sum_k \Psi^k$ converges. Just like the proof in appendix D.3, we can restrict our attention to an arbitrary mixture of (finitely many) TD models since, by proposition 1, the pass-through matrix of a canonical menu cost model is simply equivalent to a convex combination of TD pass-through matrices. For simplicity, we will start with a single TD model, and then show in step 5 of the proof how the argument extends to a mixture.

The proof in this section follows five steps. We give an overview before going into details:

1. We start with a TD pass-through matrix Ψ. We show that we can extend it to be $\mathbb{Z} \times \mathbb{Z}$ rather than $\mathbb{N}_0 \times \mathbb{N}_0$, by filling in all other elements with zeros. We show that Ψ is strictly smaller than a "shift-invariant" version of it, Ψ^{SI}.
2. The shift invariant matrix Ψ^{SI} can be fully characterized by a *symbol* function $\psi(z)$, which is a z-transform of the elements of Ψ^{SI}. We characterize $\psi(z)$.
3. We use $\psi(z)$ to characterize the infinite sum $\Psi^{SI} + \left(\Psi^{SI}\right)^2 + \left(\Psi^{SI}\right)^3 + \ldots$ as well as its first difference.

[A-11] A version of this proof also applies if only $\Phi_1 > 0$ but later survival rates Φ_t drop to zero. All that needs to be satisfied is that for any two indices $i, j \in \mathbb{N}$, there exists a power n of Ψ such that $(\Psi^n)_{i,j} > 0$. See Kemeny et al. (1976) for general conditions under which Markov chain transition matrices are transient.

4. We show that this can be used to characterize the infinite sum $\Psi + \Psi^2 + \Psi^3 + \ldots$ and its first difference as well. This will prove proposition 3 for the case of the generalized Phillips curves of TD models.

5. We prove that these results also hold for arbitrary convex combinations of finitely many TD models, and hence also hold for canonical menu cost models.

Step 1: extending Ψ to $\mathbb{Z} \times \mathbb{Z}$. With some abuse of notation, we use, from now on, the notation that Ψ refers to a $\mathbb{Z} \times \mathbb{Z}$ matrix, defined by $\Psi_{i,j}$ being as usual for $i, j \in \mathbb{N}_0$ and zero otherwise. Based on (12) we can write

$$\Psi = \underbrace{\frac{1}{\sum_{s\geq 0}\Phi_s}\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & \Phi_0 & 0 & 0 & \cdots \\ \cdots & 0 & \Phi_1 & \Phi_0 & 0 & \cdots \\ \cdots & 0 & \Phi_2 & \Phi_1 & \Phi_0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{\mathbf{A}} \cdot \underbrace{\frac{1}{\sum_{s\geq 0}\beta^s\Phi_s}\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & \Phi_0 & \beta\Phi_1 & \beta^2\Phi_2 & \cdots \\ \cdots & 0 & 0 & \Phi_0 & \beta\Phi_1 & \cdots \\ \cdots & 0 & 0 & 0 & \Phi_0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{\mathbf{B}}$$

for some survival function Φ. We define the shift-invariant version of Ψ by

$$\Psi^{SI} = \underbrace{\frac{1}{\sum_{s\geq 0}\Phi_s}\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \Phi_0 & 0 & 0 & 0 & \cdots \\ \cdots & \Phi_1 & \Phi_0 & 0 & 0 & \cdots \\ \cdots & \Phi_2 & \Phi_1 & \Phi_0 & 0 & \cdots \\ \cdots & \Phi_3 & \Phi_2 & \Phi_1 & \Phi_0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{\mathbf{A}^{SI}} \cdot \underbrace{\frac{1}{\sum_{s\geq 0}\beta^s\Phi_s}\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \Phi_0 & \beta\Phi_1 & \beta^2\Phi_2 & \beta^3\Phi_3 & \cdots \\ \cdots & 0 & \Phi_0 & \beta\Phi_1 & \beta^2\Phi_2 & \cdots \\ \cdots & 0 & 0 & \Phi_0 & \beta\Phi_1 & \cdots \\ \cdots & 0 & 0 & 0 & \Phi_0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{\mathbf{B}^{SI}}$$

Note that $\Psi^{SI} \geq \Psi$. Also note that all three, $\mathbf{A}^{SI}$, $\mathbf{B}^{SI}$, and Ψ^{SI} are *shift-invariant* (or two-sided infinite Toeplitz matrices): all entries along a diagonal are the same. This means that $\Psi^{SI}_{i+j,i}$ only depends on j and is independent of i. Denote this number by ψ_j, for $j \in \mathbb{Z}$. Similarly, define $a_j \equiv A^{SI}_{i+j,i}$ and $b_j \equiv B^{SI}_{i+j,i}$. We denote the infinite sum of powers of Ψ by $\mathbf{N} = \Psi + \Psi^2 + \ldots$ and its shift-invariant version by

$$\mathbf{N}^{SI} \equiv \Psi^{SI} + \left(\Psi^{SI}\right)^2 + \left(\Psi^{SI}\right)^3 + \ldots$$

Step 2: the symbol function $\psi(z)$ of Ψ^{SI}. Defining the z-transforms $a(z) \equiv \sum_{j=0}^{\infty} a_j z^j$, $b(z) \equiv \sum_{j=-\infty}^{0} b_j z^j$, $\psi(z) \equiv \sum_{j=-\infty}^{\infty} \psi_j z^j$, and $\Phi(z) \equiv \sum_{j=0}^{\infty} \Phi_j z^j$, we can write

$$\psi(z) = a(z)b(z) = \frac{\Phi(z)\Phi(\beta z^{-1})}{\Phi(1)\Phi(\beta)} \tag{A.53}$$

where we use the fact that the product of two infinite shift-invariant matrices is given by convolution of their coefficients, and that this convolution becomes multiplication when applying the z-transform.

We recall that by our regularity assumption on TD models (see footnote 34), $\Phi(z)$ has a radius of convergence of at least $v > 1$. Hence $\Phi(\beta z^{-1})$ is analytic for $|z| > \beta v^{-1}$, and $\psi(z)$ is analytic on the annulus $\beta v^{-1} < |z| < v$.

Note also that $\psi(z)$ is strictly convex on $(\beta v^{-1}, v)$, since its Laurent expansion has all positive coefficients (being the product of $\Phi(\beta z^{-1})/\Phi(\beta)$ and $\Phi(z)/\Phi(1)$). It also satisfies $\psi(\beta) = \psi(1) = 1$, and it follows that $\psi(z) < 1$ for all $z \in (\beta, 1)$, and that β and 1 are simple zeros of $\psi(z) - 1$ because $\psi'(\beta)$ and $\psi'(1)$ are nonzero (strictly negative and positive, respectively). Further, for all non-real z such that $\beta \leq |z| \leq 1$, we must have $\psi(z) \neq 1$, since $\psi(z)$ being real for z complex implies that the triangle inequality holds strictly[A-12], $|\psi(z)| < \psi(|z|) \leq 1$.

Next, we argue that there is some $\gamma > 1$ such that on the annulus $\beta\gamma^{-1} < |z| < \gamma$, $z = \beta$ and $z = 1$ are the only zeros of $\psi(z) - 1$. We have already shown this for $\beta \leq |z| \leq 1$. Suppose to the contrary that there is no such γ, and that there exist zeros for $|z| > 1$ arbitrarily close to 1 or $|z| < \beta$ arbitrarily close to β. These zeros z must have limit points on the circles $|z| = 1$ or $|z| = \beta$, respectively, both of which are impossible since ψ is analytic and not identically zero on the annulus $\beta v^{-1} < |z| < v$.

We conclude that $\psi(z) - 1$ is analytic and has two simple zeros on some annulus $\beta\gamma^{-1} < |z| < \gamma$, with the zeros at $z = \beta$ and $z = 1$.

Step 3: characterizing $\mathbf{N}^{SI}$ and its first difference. Since the product of shift-invariant matrices on the integers is given by convolution, and this convolution becomes multiplication when applying the z-transform, it follows that $[(\Psi^{SI})^n]_{i+j,i}$, is equal to the jth coefficient of $\psi(z)^n$. We know from above that $\psi(z)$ has two simple zeros at β and 1 and is strictly less than 1 in the annulus $\beta < |z| < 1$. Hence, picking any z_l, z_h satisfying $\beta < z_l < z_h < 1$, we have $|\psi(z)| \leq M < 1$ for $z_l \leq |z| \leq z_h$. On this closed annulus, $\psi(z) + \psi(z)^2 + \ldots$ therefore converges uniformly to $\frac{\psi(z)}{1-\psi(z)}$. Hence, $[\mathbf{N}^{SI}]_{i+j,i} \equiv [\Psi^{SI} + (\Psi^{SI})^2 + \ldots]_{i+j,i}$ is given by the jth coefficient of the Laurent series of $n(z) \equiv \frac{\psi(z)}{1-\psi(z)} = \sum_{j=-\infty}^{\infty} n_j z^j$ in this region. It follows that $\mathbf{N}^{SI}$ is finite, and hence $\mathbf{N} \leq \mathbf{N}^{SI}$ is as well. This is our second proof that the infinite sum $\Psi + \Psi^2 + \ldots$ in (15) converges.

Suppose that we are interested in the first difference of the entries of $\mathbf{N}^{SI}$, i.e. $[\mathbf{N}^{SI}]_{i+j,i} -$

[A-12] More explicitly: if $\psi(z) = 1$ and z is complex, then $\psi(z) = |\mathrm{Re}\psi(z)| \leq \sum_{j=-\infty}^{\infty} \psi_j |\mathrm{Re} z^j| < \sum_{j=-\infty}^{\infty} \psi_j |z|^j = \psi(|z|)$, where the final strict inequality holds because we know that, for instance, $\psi_1 > 0$ (which follows from $\Phi_1 > 0$), and for that term $\psi_1 |\mathrm{Re} z| < \psi_1 |z|$ whenever z is not real.

$[\mathbf{N}^{SI}]_{i+j-1,i}$. This equals $n_j - n_{j-1}$, which will be the jth coefficient of the Laurent series

$$k(z) \equiv (1-z)n(z) = (1-z)\frac{\psi(z)}{1-\psi(z)} = \sum_{j=-\infty}^{\infty} k_j z^j$$

Since $\frac{\psi(z)}{1-\psi(z)}$ has a simple pole at $z = 1$ (corresponding to the simple zero of $1 - \psi(z)$), multiplying by $(1 - z)$ removes this pole, and $k(z)$ is therefore meromorphic on the annulus $\beta\gamma^{-1} < |z| < \gamma$, with the only singularity being a simple pole at $z = \beta$.

It immediately follows that $\limsup_{j\to\infty} |k_j|^{1/j} \leq \gamma^{-1} < 1$, i.e. that asymptotically as $j \to \infty$ the coefficients k_j are bounded above by some decaying exponential function. (Since the coefficients n_j are the cumulative sums of k_j, this has the useful additional implication that $\bar{n}_j$ are bounded as $j \to \infty$ as well.) Similarly, it follows that $\limsup_{j\to-\infty} |k_j|^{1/j} = \beta$.

Now consider multiplying $k(z)$ by $(\beta - z)$, to get

$$(\beta - z)k(z) \equiv \sum_{j=-\infty}^{\infty} (\beta k_j - k_{j-1})z^j \tag{A.54}$$

This removes the simple pole at $z = \beta$, and hence (A.54) is analytic on the annulus $\beta\gamma^{-1} < |z| < \gamma$. It follows that $\limsup_{j\to-\infty} |\beta k_j - k_{j-1}|^{1/j} = \beta\gamma^{-1}$, so that there exists some $M > 0$ and $n < 0$ such that $|\beta k_j - k_{j-1}| < M\beta^{-j}\gamma^j$ for all $j < n$. Extending this inequality, we note that

$$\begin{aligned}
|k_j - \beta^{-l}k_{j-l}| &\leq |k_j - \beta^{-1}k_{j-1}| + \ldots + |\beta^{-l+1}k_{j-l+1} - \beta^{-l}k_{j-l}| \\
&\leq \beta^{-1}M\beta^{-j}\gamma^j + \ldots + \beta^{-l}M\beta^{-j+l-1}\gamma^{j-l} \\
&= \beta^{-j-1}\gamma^j M\left(1 + \beta^{-1}\gamma^1 + \ldots + \beta^l\gamma^{-l}\right) \\
&< \beta^{-j-1}\gamma^j\frac{M}{1-\beta^{-1}\gamma}
\end{aligned}$$

and hence that

$$\lim_{j\to-\infty}\sup_{l\geq 0}|\beta^j k_j - \beta^{j-l}k_{j-l}| \leq \lim_{j\to-\infty}\beta^{-1}\gamma^j\frac{M}{1-\beta^{-1}\gamma} = 0 \tag{A.55}$$

i.e. that $\{\beta^j k_j\}$ is a Cauchy sequence as $j \to -\infty$. It therefore converges to some limit $\lim_{j\to-\infty}\beta^j k_j = c$. The (weaker) statement that $\lim_{j\to-\infty}\frac{k_{j-1}}{k_j} = \beta$ also immediately follows.

Step 4: using this to characterize the generalized Phillips curve K**.** Above, we have already characterized $\mathbf{N}^{SI}$ and its first difference (in rows). Our goal is now to prove that asymptotically, $\mathbf{N}^{SI}$ and $\mathbf{N}$ coincide, in the sense that for any j,

$$\lim_{i\to\infty}[\mathbf{N}^{SI} - \mathbf{N}]_{i+j,i} = 0 \tag{A.56}$$

To prove this, first we derive an expression for $\mathbf{N}^{SI} - \mathbf{N}$, writing

$$\begin{aligned}\Psi^{SI}(\mathbf{I} - \Psi) - (\mathbf{I} - \Psi^{SI})\Psi &= \Psi^{SI} - \Psi \\ (\mathbf{I} - \Psi^{SI})^{-1}\Psi^{SI} - \Psi(\mathbf{I} - \Psi)^{-1} &= (\mathbf{I} - \Psi^{SI})^{-1}(\Psi^{SI} - \Psi)(\mathbf{I} - \Psi)^{-1} \\ \mathbf{N}^{SI} - \mathbf{N} &= \tilde{\mathbf{N}}^{SI}(\Psi^{SI} - \Psi)\tilde{\mathbf{N}} \end{aligned} \tag{A.57}$$

where in the last line we used $\mathbf{N} = \Psi + \Psi^2 + \ldots = \Psi(\mathbf{I} - \Psi)^{-1}$ (and the same for $\mathbf{N}^{SI}$), and defined $\tilde{\mathbf{N}} \equiv \mathbf{I} + \mathbf{N} = (\mathbf{I} - \Psi)^{-1}$ (and the same for $\tilde{\mathbf{N}}^{SI}$).

Let us first characterize the matrix in the middle on the right of (A.57), $\Psi^{SI} - \Psi$. We recall that $\Psi^{SI} = \mathbf{A}^{SI}\mathbf{B}^{SI}$ and $\Psi = \mathbf{AB}$, and note that actually also $\Psi = \mathbf{A}^{SI}\mathbf{B}$, since $\mathbf{A}^{SI}$ and $\mathbf{A}$ only differ in columns $j < 0$, but all rows $i < 0$ in $\mathbf{B}$ are zero. Hence $\Psi^{SI} - \Psi = \mathbf{A}^{SI}\left(\mathbf{B}^{SI} - \mathbf{B}\right)$.

We next observe that

$$\sum_i (\Psi^{SI}_{ij} - \Psi_{ij}) = \sum_i (B^{SI}_{ij} - B_{ij}) = \frac{\sum_{r=j+1}^{\infty} \beta^r \Phi_r}{\sum_{r=0}^{\infty} \beta^r \Phi_r} \tag{A.58}$$

where the first equality follows because $\mathbf{A}^{SI}$ is column-stochastic, and the second equality follows directly from the definitions of $\mathbf{B}^{SI}$ and $\mathbf{B}$. We observe that for some sufficiently large C and all $j \geq 0$, (A.58) is bounded above by $C\beta^j$.

From our earlier characterization, we know that Ψ^{SI} and therefore $\tilde{\mathbf{N}}^{SI} = \mathbf{I} + \mathbf{N}^{SI}$ has all entries bounded above by some M. It follows that all entries in the jth column of $\tilde{\mathbf{N}}^{SI}(\Psi^{SI} - \Psi)$ are bounded by $MC\beta^j$. Using (A.57), we conclude that

$$\begin{aligned}\lim_{i\to\infty}[\mathbf{N}^{SI} - \mathbf{N}]_{i+j,i} &= \lim_{i\to\infty}[\tilde{\mathbf{N}}^{SI}(\Psi^{SI} - \Psi)\tilde{\mathbf{N}}]_{i+j,i} \\ &\leq \lim_{i\to\infty}[\tilde{\mathbf{N}}^{SI}(\Psi^{SI} - \Psi)\tilde{\mathbf{N}}^{SI}]_{i+j,i} \\ &\leq \lim_{i\to\infty}\sum_{k=0}^{\infty} MC\beta^k \tilde{\mathbf{N}}^{SI}_{k,i} \\ &= MC\sum_{k=0}^{\infty}\beta^k \lim_{i\to\infty}\tilde{\mathbf{N}}^{SI}_{k,i} = 0\end{aligned}$$

which, since $\mathbf{N}^{SI} \geq \mathbf{N}$, implies that $\lim_{i\to\infty}[\mathbf{N}^{SI} - \mathbf{N}]_{i+j,i} = 0$. It follows that

$$\lim_{i\to\infty}(\mathbf{N}^{SI}_{i+j,i} - \mathbf{N}^{SI}_{i+j-1,i}) - (\mathbf{N}_{i+j,i} - \mathbf{N}_{i+j-1,i}) = 0$$

as well. Since the generalized Phillips curve $\mathbf{K} = (I - L)(\Psi + \Psi^2 + \ldots)$ has entries equal to $\mathbf{K}_{i+j,i} = \mathbf{N}_{i+j,i} - \mathbf{N}_{i+j-1,i}$, it follows that its columns asymptotically approach the same two-sided sequence around the diagonal as in the rows of $\mathbf{N}^{SI}_{i+j,i} - \mathbf{N}^{SI}_{i+j-1,i}$, which we already characterized in the previous step as the sequence $\{k_j\}$.

Step 5: extending to a mixture of multiple time-dependent models. Suppose we have a mixture of multiple time-dependent models $\ell = 1, \ldots, n$, each with its own survival function Φ^ℓ and pass-through matrix Ψ^ℓ, with weights c_ℓ summing to 1. This mixture will have pass-through matrix $\Psi = c_1\Psi^1 + \ldots + c_n\Psi^n$.

We can construct $\Psi^{SI,\ell}$ as before for each Ψ^ℓ, and combine to obtain $\Psi^{SI} = c_1\Psi^{SI,1} + \ldots + c_n\Psi^{SI,n}$, which is still shift-invariant. We then obtain the same characterization of the z-transform $\psi(z)$ of Ψ^{SI} as before. In particular, since $\psi(z)$ is a mixture of the underlying ℓ, it is still analytic in some annulus $\beta v^{-1} < |z| < v$ for $v > 1$ (where we can take the minimum v^ℓ across all ℓ) and remains convex on $(\beta v^{-1}, v)$ with simple zeros at β and 1. It follows from our arguments in step 2 that $\psi(z) - 1$ is analytic, strictly smaller than 1 for $\beta < |z| < 1$, and has two simple zeros $z = \beta$ and $z = 1$ on some annulus $\beta\gamma^{-1} < |z| < \gamma$. Given these properties of $\psi(z)$, step 3 is unchanged.

For step 4, the identity (A.57) remains unchanged, and we can use the argument from (A.58) to show that for each ℓ, the jth column of $\Psi^{SI,\ell} - \Psi^\ell$ is bounded above by $C^\ell\beta^j$ for some constant C^ℓ. Taking $C \equiv \max_\ell C^\ell$, it follows that the jth column of $\Psi^{SI} - \Psi$ is bounded above by $C\beta^j$, and the rest of the proof goes through as before, concluding our argument.

D.5 Robustness of the numerical equivalence result

This appendix provides robustness exercises for the numerical equivalence result from section 4. We show that the approximate numerical equivalence holds for several extensions of the baseline menu cost model used in the main text.

D.5.1 Trend inflation

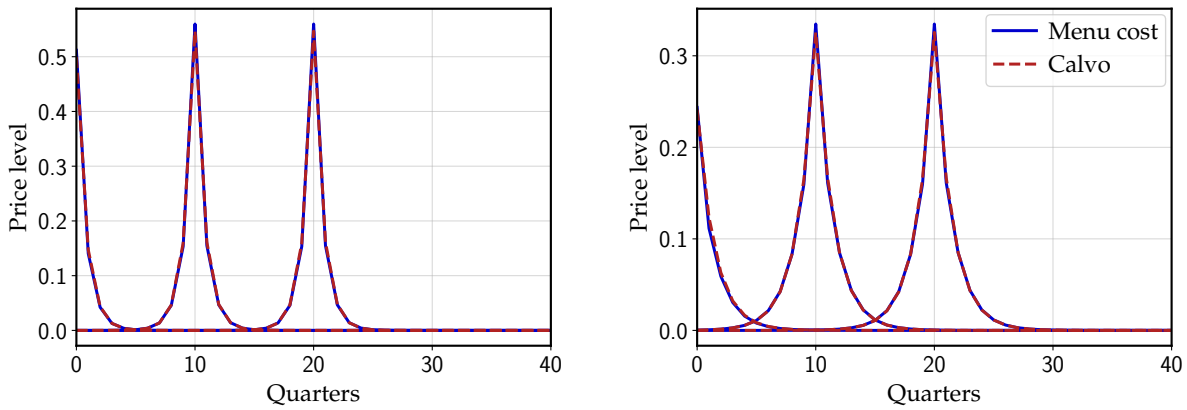
First, we introduce trend inflation. This can be done by adding a drift term $\mu > 0$ to the law of motion for the static optimal price (1), which becomes:

$$p^*_{it} = p^*_{it-1} + \mu + \epsilon_{it}.$$

Since the time unit is one quarter, the drift μ corresponds to an annual inflation rate of 4μ. We compute pass-through matrices and generalized Phillips curves in this case using proposition 7.

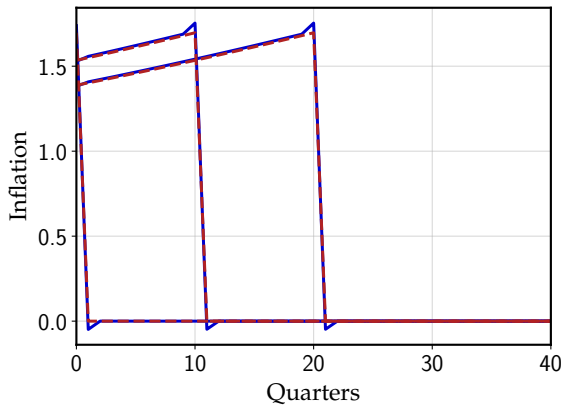
Figure D.1 shows pass-through and Phillips curve matrices for both the GL and NS models with annual inflation rate of 2%, while figure D.2 does the analogous exercise for a 5% annual inflation rate. The state-dependent pass-through matrices are still indistinguishable from the corresponding Calvo approximations. For the Phillips curve matrices, it is visible that the Calvo approximation is slightly better for lower inflation, although the fit is still very good for moderate inflation levels.

The upper graphs of Figure D.3, panels (a) and (b), show in solid blue the best-fitting κ when we calibrate the baseline GL and NS models, and then vary the level of trend inflation μ. The figures show that the slope sharply increases with trend inflation, consistent with the fact that trend inflation increases the steady-state frequency of price adjustment (see, e.g., Alvarez, Beraja,

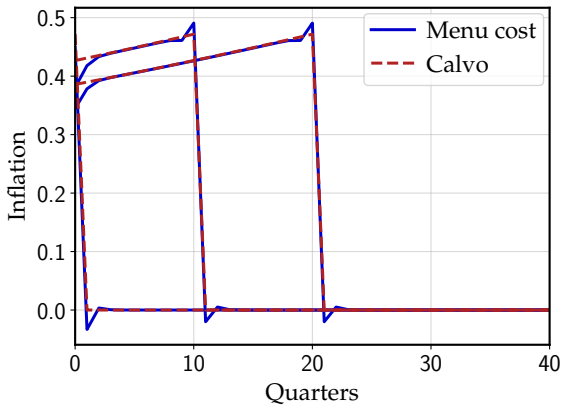


(a) Golosov-Lucas pass-through matrix

(b) Nakamura-Steinsson pass-through matrix



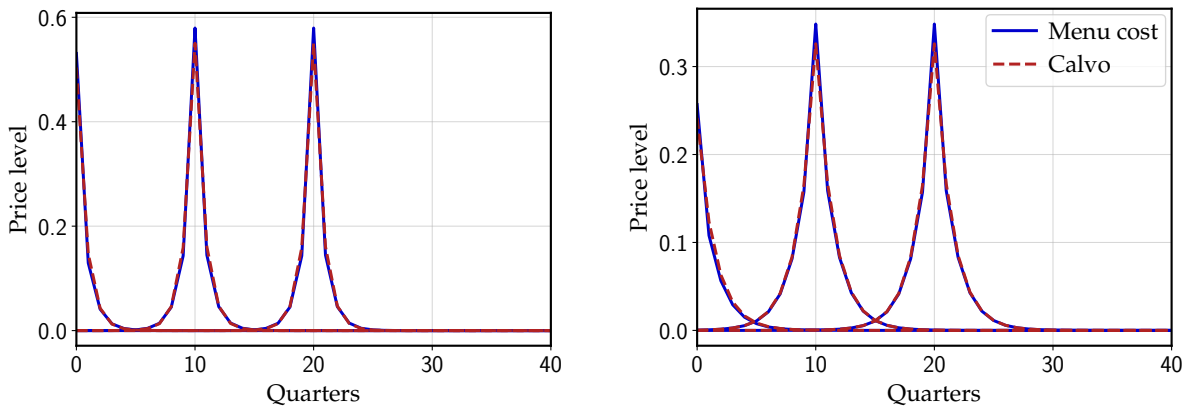
(c) Golosov-Lucas generalized Phillips curve



(d) Nakamura-Steinsson generalized Phillips curve

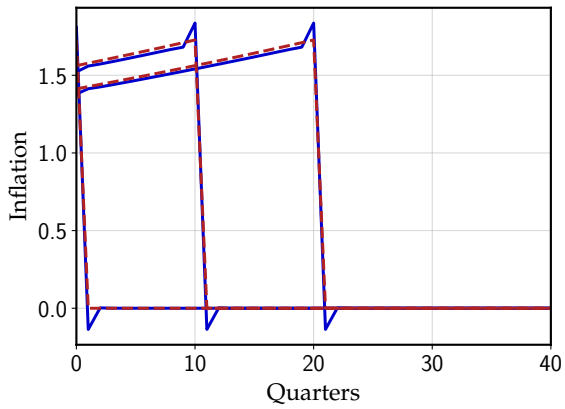
Figure D.1: Menu cost models and Calvo approximations with 2% annual trend inflation

Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the GL and NS models, calibrated to match the same empirical moments as in the main text, as well as the best-fitting Calvo approximations.

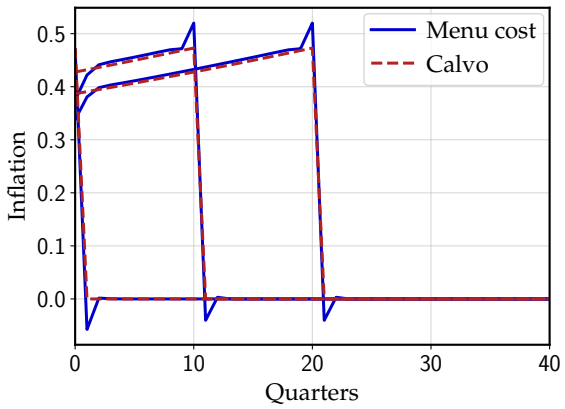


(a) Golosov-Lucas pass-through matrix

(b) Nakamura-Steinsson pass-through matrix



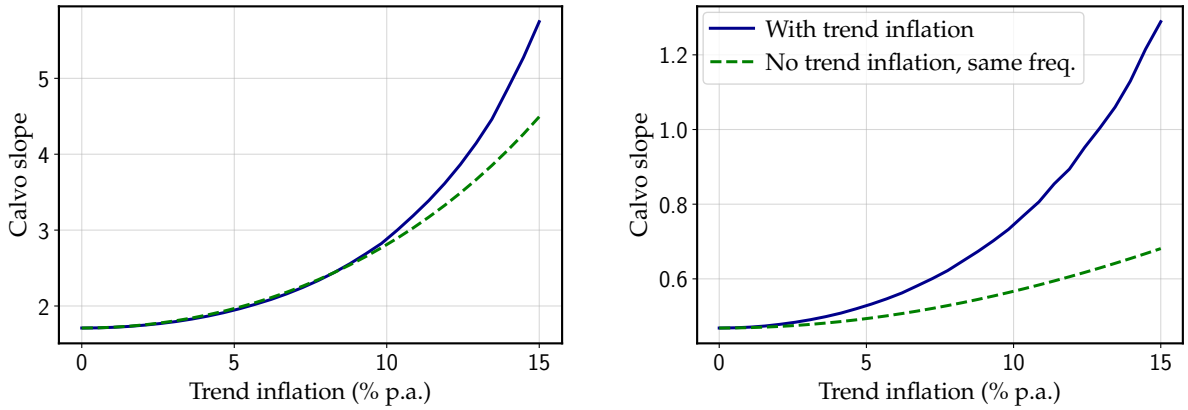
(c) Golosov-Lucas generalized Phillips curve



(d) Nakamura-Steinsson generalized Phillips curve

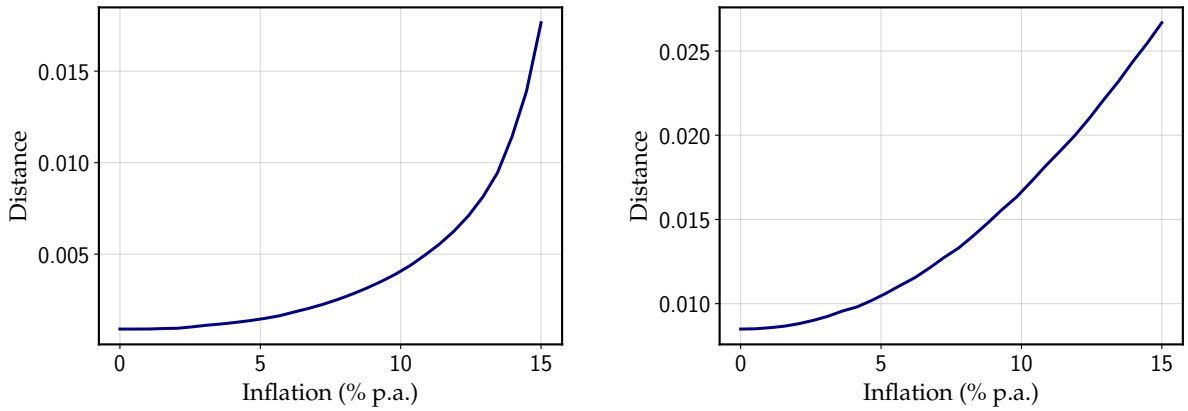
Figure D.2: Menu cost models and Calvo approximations with 5% annual trend inflation

Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the GL and NS models, calibrated to match the same empirical moments as in the main text, as well as the best-fitting Calvo approximations.



(a) Golosov-Lucas – approx. Calvo slopes

(b) Nakamura-Steinsson – approx. Calvo slopes



(c) Golosov-Lucas – approx. distance

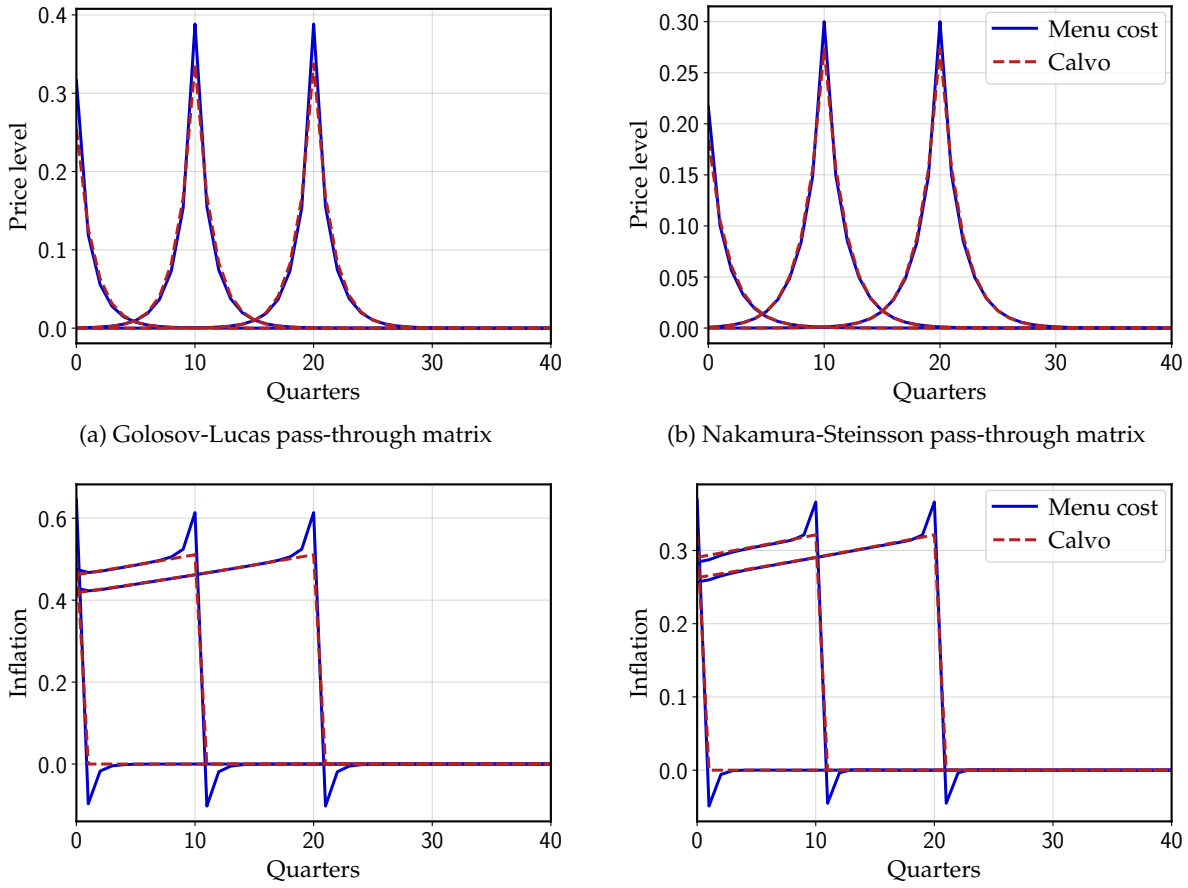
(d) Nakamura-Steinsson – approx. distance

Figure D.3: Best-fitting Calvo Phillips curve slope and distance for various levels of trend inflation μ

Note: The top two panels are constructed as follows. For each model calibrated as in 2.5, we vary trend inflation and compute the resulting adjustment frequency f and best-fitting Calvo slope κ, shown in the solid blue line. We then compute the Calvo approximation of a model *without* trend inflation, but with frequency f, shown in the dashed green line. The bottom two panels show the adjustment distance from approximating the models *with* trend inflation by a Calvo model as a function of trend inflation.

Gonzalez-Rozada and Neumeyer 2019). The dashed green lines show the κ that approximates an equivalent model *without* trend inflation, but with an adjustment frequency calibrated to match the one from the corresponding model *with* trend inflation. We see that both lines diverge, most noticeably for the Nakamura-Steinsson model, which suggests that trend inflation increases price flexibility not only by increasing the adjustment frequency, but also by magnifying the selection effect.

The bottom graphs, panels (c) and (d), show the distance between the SD model and the Calvo approximation, which is increasing with inflation. However, the distance increases very modestly for low inflation levels, which is intuitive. Adding positive or negative inflation changes the pass-through matrix and the generalized Phillips curve in the same way. As a consequence of this symmetry, the distance lines have zero slope at $\mu = 0$.



(a) Golosov-Lucas pass-through matrix

(b) Nakamura-Steinsson pass-through matrix

(c) Golosov-Lucas generalized Phillips curve

(d) Nakamura-Steinsson generalized Phillips curve

Figure D.4: Menu cost models and Calvo approximations with infrequent shocks ($p = 0.5$)

Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the GL and NS models, calibrated to match the same empirical moments as in the main text, as well as the best-fitting Calvo approximations.

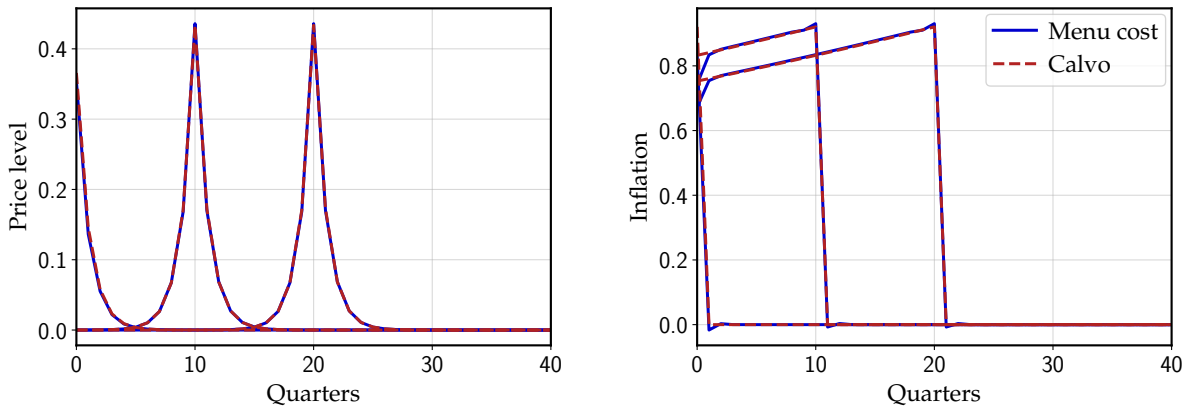
D.5.2 Infrequent shocks

Now we return to the model with no trend inflation, and instead introduce infrequent shocks, as in Midrigan (2011). More specifically we assume that idiosyncratic shocks follow

$$\epsilon_{it} = \begin{cases} 0 & \text{with probability } 1-p \\ N(0, \sigma_\epsilon^2) & \text{with probability } p \end{cases}.$$

This effectively increases the kurtosis of the (unconditional) shock distribution, so this feature is often referred to in the literature as leptokurtic shocks.[A-13] Figure D.4 shows results. Similarly to the trend inflation case, the pass-through matrices of the menu cost models are still indistinguishable from their Calvo approximations. The fit of the generalized Phillips curve slightly deteriorates, but is still very good.

[A-13] An alternative approach in Karadi and Reiff (2019) assumes a mixture of two normal distributions.



(a) Pass-through matrix

(b) Generalized Phillips curve

Figure D.5: Multi-product model and Calvo approximations

Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the multi-product model, calibrated to match the moments in table 1, as well as the best-fitting Calvo approximation.

D.5.3 Multi-product model

The next extension we explore is a multi-product model, as in Midrigan (2011) and Alvarez and Lippi (2014). We revert to our baseline model, without trend inflation or leptokurtic shocks, and now assume each firm sells two distinct products. The state variable of the firm optimization problem is now a pair of price gaps $(x_{it,1}, x_{it,2})$, each one evolving independently as a random walk without drift, although both are subject to the same aggregate marginal cost shock. The loss function is now given by

$$\frac{1}{2}\left(x_{it,1} - \log MC_t\right)^2 + \frac{1}{2}\left(x_{it,2} - \log MC_t\right)^2.$$

Importantly, firms face economies of scope in price adjustments – there is a single menu cost whose payment allows the firm to adjust the prices of both its products. Otherwise, aggregate dynamics would be the same as in a single-product model. Figure D.5 shows results for this case. Again, the Calvo approximation is very precise.

D.5.4 Multi-sector model

The next extension we analyze is a multi-sector economy. Consider an economy composed of N economic sectors, each one characterized by its own parameter values, i.e., potentially different menu costs, probabilities of free adjustments, and volatility of idiosyncratic shocks. Each sector, indexed by $j \in \{1, \ldots, N\}$, has weight ω_j in the price index, in such a way that the log aggregate price level is

$$p_t = \sum_{j=1}^{N} \omega_j p_{jt},$$

Sector	Weight (%)	Frequency (%)	Abs. size (%)
Vehicle fuel, used cars	7.7	91.6	4.9
Utilities	5.3	49.4	6.4
Travel	5.5	43.7	18.4
Unprocessed food	5.9	25.4	15.9
Transport goods	8.3	21.3	8.9
Services (1)	7.7	21.7	4.0
Processed food, other goods	13.7	11.9	11.4
Services (2)	7.5	8.4	6.7
Household furnishings	5.0	6.5	10.1
Services (3)	7.8	6.2	8.8
Rec. goods	3.6	6.1	10.2
Services (4)	7.6	4.9	8.1
Apparel	6.5	3.6	12.4
Services (5)	7.9	2.9	13.5

Table D.2: Sectoral pricing moments.

Note: this table reproduces the data from Nakamura and Steinsson (2010), and shows CPI weights, monthly frequency and mean absolute size of price changes for each sector. Services are sorted into five groups according to entry-level adjustment frequency in the CPI. See Nakamura and Steinsson (2010) for more details. For our calibration, we convert monthly frequencies $f_{monthly}$ into quarterly frequencies $f_{quarterly} = 1 - (1 - f_{monthly})^3$.

where p_{jt} is the sectoral price level of sector j. From the above equation, it follows that the pass-through matrix of the multi-sector economy Ψ is given by the same weighted average of the sectoral pass-through matrices Ψ_j:

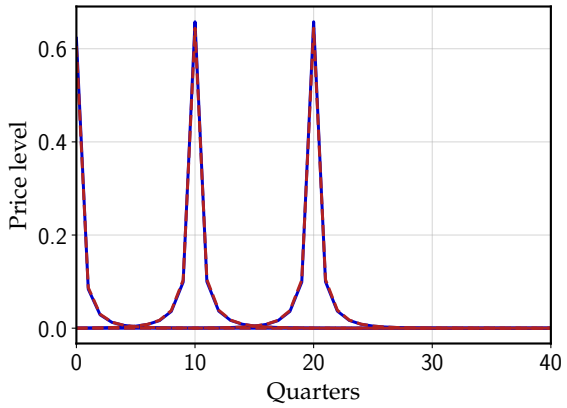
$$\Psi = \sum_{j=0}^{J} \omega_j \Psi_j.$$

Once we have this pass-through matrix, we can apply the same transformation in (16) to obtain the generalized Phillips curve. Consequently, if each Ψ_k can be well approximated by a Calvo model, then the multi-sector state-dependent economy will be close to a multi-sector Calvo one.

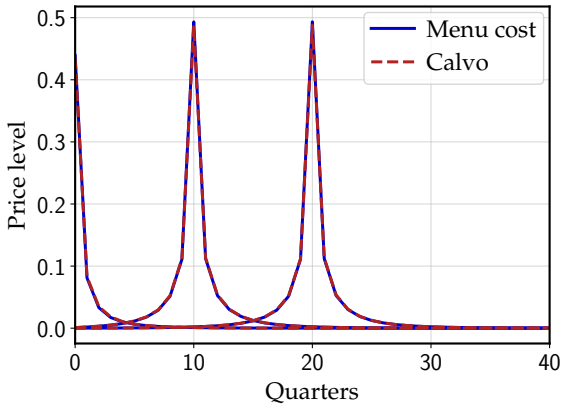
Following the approach outlined above, we calibrate a 14-sector menu cost economy using the same moments as Nakamura and Steinsson (2010), reproduced in table D.2. For each sector, we find the best-fitting Calvo model and compute the corresponding aggregate pass-through and Phillips curve matrices. Figure D.6 shows results. For both our main specifications – with and without free adjustments –, the two models are again almost identical.

D.5.5 Large shocks

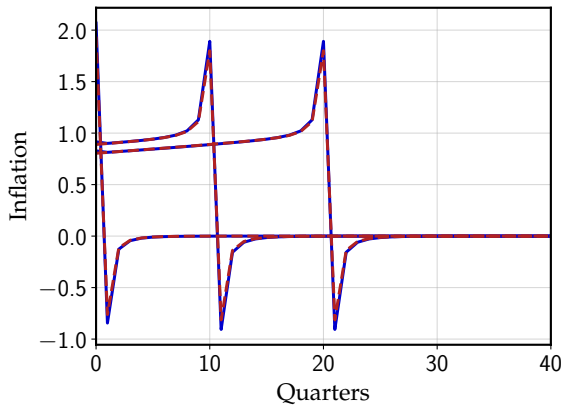
Finally, we study how well the Calvo approximation fares in comparison to large, nonlinear marginal cost shocks in state-dependent models. State-dependent models are well-known for featuring nonlinearities: a large aggregate shock may endogenously trigger many price adjustments, increasing the flexibility of the aggregate price level in response to it. In order to assess this effect,



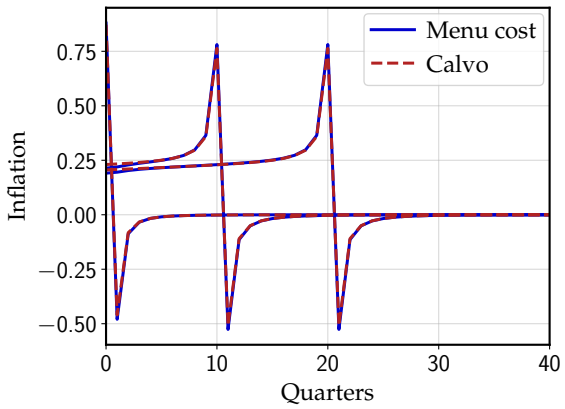
(a) Multi-sector Golosov-Lucas pass-through matrix



(b) Multi-sector Nakamura-Steinsson pass-through matrix



(c) Multi-sector Golosov-Lucas generalized Phillips curve



(d) Multi-sector Nakamura-Steinsson generalized Phillips curve

Figure D.6: Multi-sector menu cost models and Calvo approximations

Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the multi-sector GL and NS models, calibrated to match the moments in table D.2, as well as the best-fitting Calvo approximations.

we compute the nonlinear price responses to nominal marginal cost shocks of the form

$$MC_t = MC_0 \rho^t$$

for $MC_0 \in \{2.5\%, 5\%\}$. We compare the price responses to the best-fitting linear Calvo approximation. Results are shown in figure D.7. For shocks of initial size 2.5%, nonlinearities are negligible and the Calvo approximation again provides almost identical impulse responses. For a large shock of initial size 5%, we start to see some discrepancies for the Golosov-Lucas model, in which the extensive margin of adjustment is stronger. For the Nakamura-Steinsson model, on the other hand, the Calvo model still provides indistinguishable responses. For a shock of size 10%, however, the Calvo approximation breaks down for the Golosov-Lucas model, but less so for the Nakamura-Steinsson model. This is because, as we described in section 4.2, the *Ss* band of the GL model is much tighter than for the NS model. A tighter *Ss* band makes it easier for significant mass to fall outside the *Ss* band for a shift in the band of a given size.

Figure D.8 examines the extent of aggregate nonlinearity in the generalized Phillips curve for our two main calibrations of the canonical menu cost model. We conduct two exercises. In the top graphs, panels (a) and (b), starting from the steady state, we consider a large AR(1) shock to real marginal cost that increases annual inflation on impact by 5% (quarterly by 1.25%). We compare the resulting nonlinear impulse response in the menu cost model to the linear impulse response in the same model, and to the linear impulse response of the approximating Calvo model from section 4. We find that the three impulse responses all closely match, with the nonlinear menu cost slightly above the linear in the first period and very close thereafter.

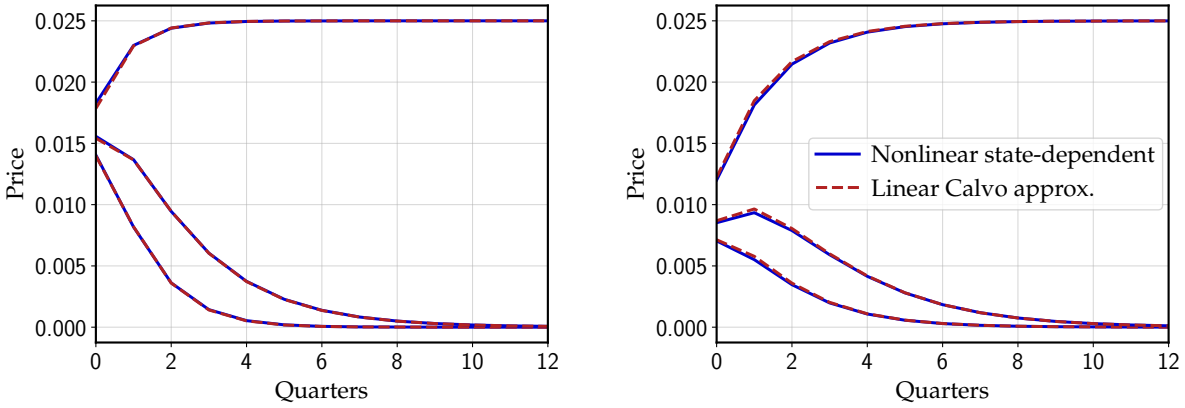
In the middle graphs, panels (c) and (d), we perform the same experiment, but for a shock that generates a 10% annualized inflation on impact for the Calvo model. Here, we see that nonlinearities start to play a significant role, especially in the GL calibration, as the NS model features a larger inaction region. Therefore, for very large real shocks there may be a noticeable discrepancy between the linear and nonlinear responses.

In the bottom graphs, panels (e) and (f), we perform a similar exercise, but assuming that a 5% shock at date 0 comes on top of a same-sized shock at date -5, so that the initial distribution of price gaps is potentially away from the steady state coming into date 0. We find limited departure between this nonlinear impulse response and the linear impulse responses, although the initial gap between nonlinear and linear menu cost is now slightly larger in the NS model, indicating only mild state dependence.

E Appendix to Section 5

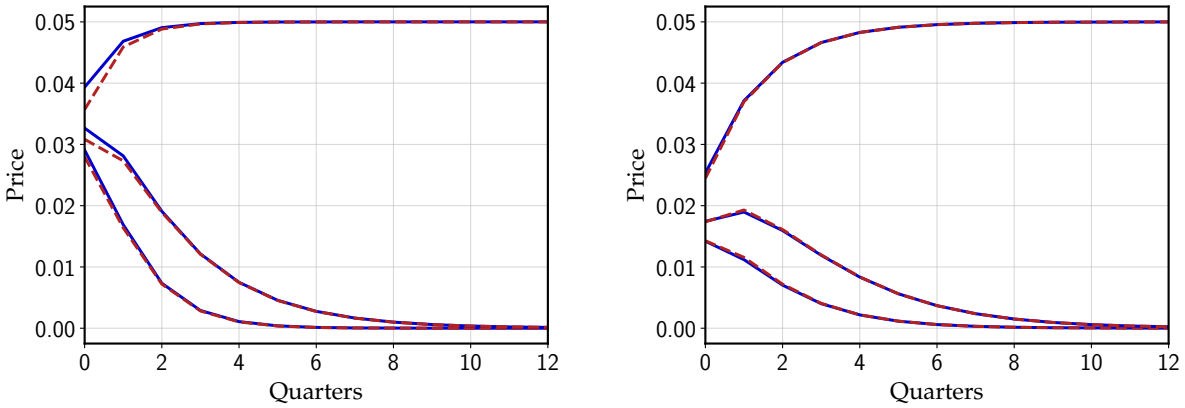
E.1 The log-linearized system for non-infinitesimal $\sigma_\epsilon > 0$

We begin by collecting all the equilibrium conditions of the model. There are 9 conditions for 9 unknown sequences, $\{w_t, W_t, P_t, Y_t, N_t, \Delta_t, \Xi_t, i_t, \pi_t\}$. The conditions are:



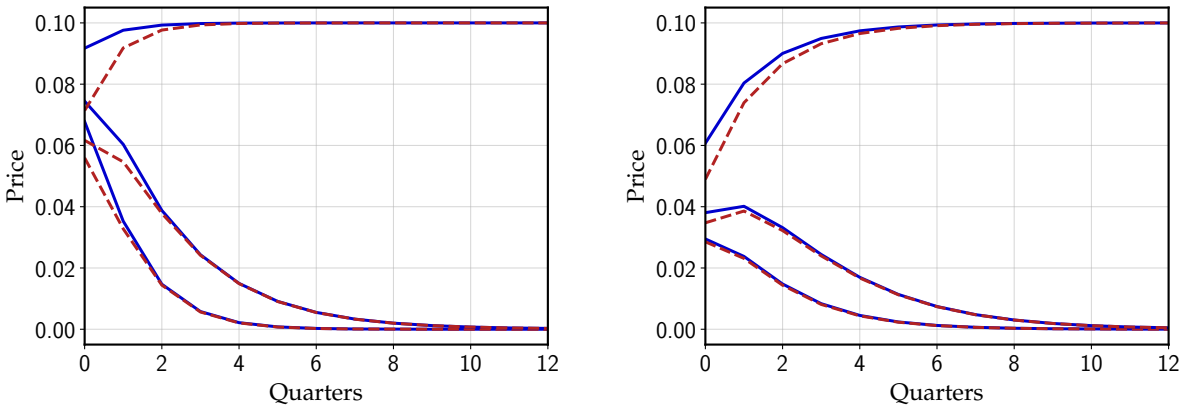
(a) Golosov-Lucas, 2.5% nominal shocks

(b) Nakamura-Steinsson, 2.5% nominal shocks



(c) Golosov-Lucas, 5% nominal shocks

(d) Nakamura-Steinsson, 5% nominal shocks

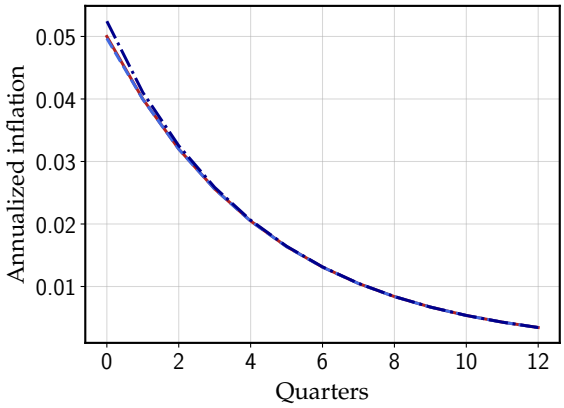


(e) Golosov-Lucas, 10% nominal shocks

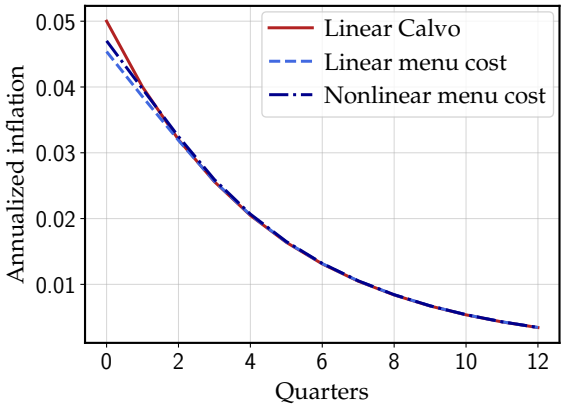
(f) Nakamura-Steinsson, 10% nominal shocks

Figure D.7: Price level responses to AR(1) nonlinear nominal marginal cost shocks

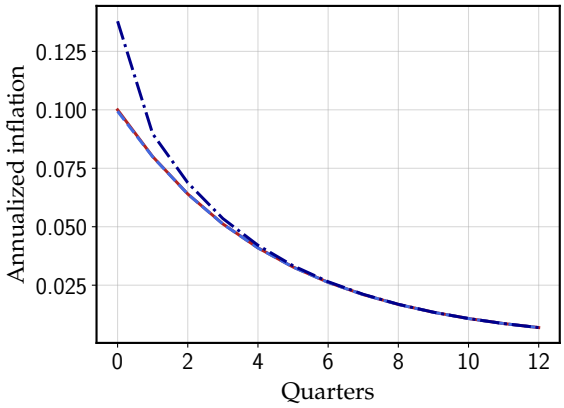
Note: nonlinear impulse responses to AR(1) marginal cost shocks for GL and NS models, calibrated as in table 1, as well as linear responses for the best-fitting Calvo approximations. Shock persistence values are $\{0.3, 0.6, 1\}$.



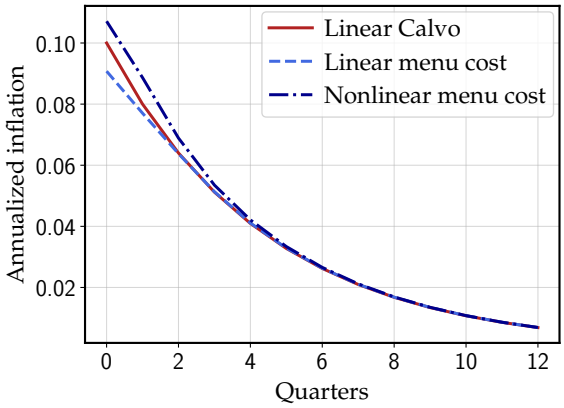
(a) Golosov-Lucas, 5% nonlinear real shock.



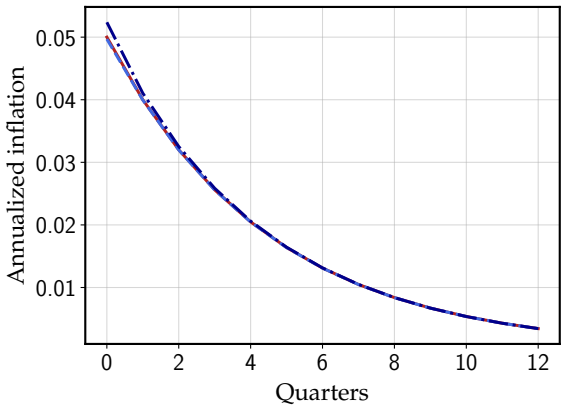
(b) Nakamura-Steinsson, 5% nonlinear real shock.



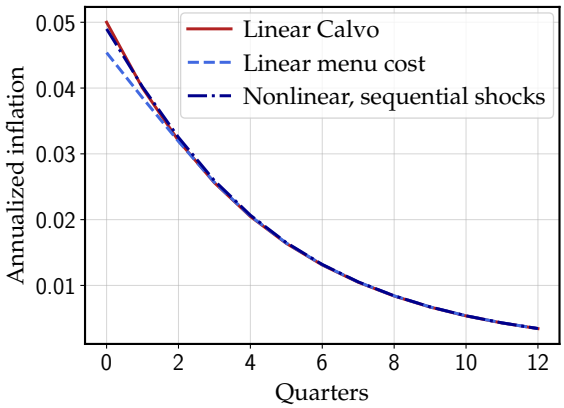
(c) Golosov-Lucas, 10% nonlinear real shock.



(d) Nakamura-Steinsson, 10% nonlinear real shock.



(e) Golosov-Lucas, 5% sequential nonlinear real shocks.



(f) Nakamura-Steinsson, 5% sequential nonlinear real shocks.

Figure D.8: Aggregate nonlinearity of the generalized Phillips curve

Note: linear and nonlinear impulse responses to AR(1) real marginal cost shocks for multi-sector GL and NS models, calibrated as in table 1, as well as linear responses for the best-fitting Calvo approximations. The shock persistence is 0.8 and the size is determined so as to generate a 5% or 10% annualized increase in inflation in the Calvo model. In panels (e) and (f), this shock comes after an identically-sized shock at date -5.

- **Optimal pricing behavior of firms:** Price setters solve (50) for given $\{W_t, w_t, Y_t\}$,

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t Y_t^{-\sigma} \left[\left(\frac{\zeta}{\zeta - 1} w_t \right)^{1-\zeta} Y_t \cdot F\left(x_{it} - \log W_t\right) + \sigma_\epsilon^2 \xi_{it} w_t 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}} \right] \tag{A.59a}$$

From the optimal price gaps x_{it} we can then compute the path of the price level, as in (49),

$$P_t = \frac{\zeta}{\zeta - 1} \left(\int_0^1 e^{(1-\zeta) x_{it}} di \right)^{\frac{1}{1-\zeta}}$$

as well as price dispersion

$$\Delta_t \equiv \left(\int_0^1 e^{(1-\zeta) x_{it}} di \right)^{\frac{\zeta}{1-\zeta}} \int_0^1 e^{-\zeta x_{it}} di \geq 1$$

and the aggregate amount of labor used for menu cost, as in (51),

$$\Xi_t \equiv \int_0^1 \sigma_\epsilon^2 \xi_{it} 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}} di$$

Together, we can summarize the pricing behavior as three sequence-space functions,

$$P_t = \mathcal{P}_t\left(\{W_s, w_s, Y_s\}\right), \quad \Delta_t = \mathcal{D}_t\left(\{W_s, w_s, Y_s\}\right), \quad \Xi_t = \mathcal{X}_t\left(\{W_s, w_s, Y_s\}\right) \tag{A.59b}$$

- **Labor demand:** Labor needs to be consistent with optimal labor demand by firms, (52),

$$N_t = Y_t \Delta_t + \Xi_t \tag{A.59c}$$

- **Labor supply:** Labor needs to be consistent with optimal labor supply by households, (42)

$$b N_t^{\varphi} = w_t Y_t^{-\sigma} \tag{A.59d}$$

where we already substituted out consumption using goods market clearing $C_t = Y_t$.

- **Monetary policy rule:** The nominal interest rate follows the Taylor rule

$$i_t = i_{ss} + \phi \pi_t + \nu_t \tag{A.59e}$$

- **Euler equation:** Output follows the household's Euler equation (42)

$$Y_t^{-\sigma} = \frac{P_t}{P_{t+1}} \left(1 + i_t\right) Y_{t+1}^{-\sigma} \tag{A.59f}$$

- **Real wage:** The real wage is defined as

$$w_t = W_t / P_t \tag{A.59g}$$

- **Inflation:** Inflation is given by

$$\pi_t = P_t / P_{t-1} - 1 \tag{A.59h}$$

We log-linearize equations (A.59b)–(A.59h) around a deterministic steady state in which $\nu_t = \nu_{ss} = 0$. We find

$$\hat{\mathbf{P}} = \mathbf{J}^{P,W}\hat{\mathbf{W}} + \mathbf{J}^{P,Y}\hat{\mathbf{Y}} + \mathbf{J}^{P,w}\hat{\mathbf{w}}$$

$$\hat{\mathbf{\Delta}} = \mathbf{J}^{\Delta,W}\hat{\mathbf{W}} + \mathbf{J}^{\Delta,Y}\hat{\mathbf{Y}} + \mathbf{J}^{\Delta,w}\hat{\mathbf{w}}$$

$$\hat{\mathbf{\Xi}} = \mathbf{J}^{\Xi,W}\hat{\mathbf{W}} + \mathbf{J}^{\Xi,Y}\hat{\mathbf{Y}} + \mathbf{J}^{\Xi,w}\hat{\mathbf{w}}$$

$$\hat{N}_t = \frac{Y_{ss}\Delta_{ss}}{Y_{ss}\Delta_{ss} + \Xi_{ss}}\left(\hat{Y}_t + \hat{\Delta}_t\right) + \frac{\Xi_{ss}}{Y_{ss}\Delta_{ss} + \Xi_{ss}}\hat{\Xi}_t$$

$$\varphi\hat{N}_t = \hat{w}_t - \sigma\hat{Y}_t$$

$$\hat{i}_t = \phi\hat{\pi}_t + \hat{\nu}_t$$

$$\hat{Y}_t = \hat{Y}_{t+1} - \frac{1}{\sigma}(\hat{i}_t - \hat{\pi}_{t+1})$$

$$\hat{w}_t = \hat{W}_t - \hat{P}_t$$

$$\hat{\pi}_t = \hat{P}_t - \hat{P}_{t-1}$$

Here, we denote by $\mathbf{J}^{P,X}$ the Jacobian of $\{\mathcal{P}_t\}$ with respect to sequence $\{X_s\}$, and similarly for $\mathbf{J}^{\Delta,X}$ and $\mathbf{J}^{\Xi,X}$.

E.2 Characterizing firm Jacobians in general case

The Jacobians $\mathbf{J}$ in the previous section are no longer given by an exact equivalence result of the simple form (31). Here, instead, we retrace the steps of proposition 1 to more generally characterize the Jacobian of an arbitrary aggregate outcome to an arbitrary shock, in the menu cost model where the period loss function is some arbitrary $F(\cdot)$—including, for instance, the case of non-infinitesimal $\sigma_\epsilon > 0$.

Law of motion for arbitrary aggregate. Suppose that we have some aggregate outcome $Y(g^{end})$ where g^{end} is the end-of-period density over x (and is an ordinary smooth density plus a Dirac delta at x^*). The examples that will ultimately be relevant to us are the price level $P(g^{end}) = \frac{\zeta}{\zeta-1}\left(\int e^{(1-\zeta)x}g^{end}(x)dx\right)^{\frac{1}{1-\zeta}}$, the price dispersion index $\Delta(g^{end}) = \left(\int e^{(1-\zeta)x}g^{end}(x)dx\right)^{\frac{\zeta}{1-\zeta}}\int e^{-\zeta x}g^{end}(x)dx$, and the total number of adjusters $ADJ(g^{end}) = \int \mathbf{1}_{x=0}g^{end}(x)dx$ (which determines menu costs

paid Ξ).[A-14]

Now, define $E^{Y,0}(x)$ to be the gradient of Y with respect to g^{end} around the steady state, and $E^{Y,t}(x)$ recursively as the expectation of $E^{Y,t-1}(x)$ given the steady-state policies. Following the proof of proposition 1, suppose that at date $t-s$, there is a one-time change in policies $\overline{x}_{t-s}$, $\underline{x}_{t-s}$, and x^*_{t-s}. Then we have

$$dY_t = (E^{Y,s}(\overline{x}) - E^{Y,s}(x^*))g^{end}(\overline{x})d\overline{x}_{t-s} - (E^{Y,s}(\underline{x}) - E^{Y,s}(x^*))g^{end}(\underline{x})d\underline{x}_{t-s} + \text{freq} \cdot (E^{Y,s})'(x^*)dx^*_{t-s}$$

Combining across all periods, and noting that $g^{end}(x) = (1-\lambda)g(x)$ at all points inside the adjustment bands (except that g^{end} also has a Dirac delta at x^*), we have the law of motion

$$\begin{aligned} dY_t = (1-\lambda)g(\overline{x}) \sum_{s=0}^{t} (E^{Y,s}(\overline{x}) - E^{Y,s}(x^*))d\overline{x}_{t-s} \\ - (1-\lambda)g(\underline{x}) \sum_{s=0}^{t} (E^{Y,s}(\underline{x}) - E^{Y,s}(x^*))d\underline{x}_{t-s} + \text{freq} \cdot \sum_{s=0}^{t} (E^{Y,s})'(x^*)dx^*_{t-s} \end{aligned} \quad \text{(A.60)}$$

Policy function for arbitrary input. Now suppose that the flow payoff function, excluding the menu cost, is given by some arbitrary $F(x, Z)$, where Z is any time-varying aggregate input. We define $E^{Z,0}(x)$ to be the derivative of $F(x, Z)$ with respect to Z around the aggregate steady state, and $E^{Z,t}(x)$ recursively given $E^{Z,t-1}(x)$. Similarly, we define $E^{F,0}(x)$ to be the derivative of $F(x, Z)$ with respect to x around the aggregate steady state, and $E^{F,t}(x)$ recursively.

Now, following the proof, we suppose there is a shock dZ at date s. Then we have $dV_s(x) = E^{Z,0}(x)dZ$, and by the same envelope argument, $dV_t(x) = \beta^{s-t}E^{Z,s-t}(x)dZ$. We also have $V'(x) = \sum_{u=0}^{\infty} \beta^u E^{F,u}(x)$, using the same envelope argument as in appendix C.1.

Similarly, if there is a shock $d\xi$ at date s, we have a change in total value at time s of $E^{ADJ,0}(x)d\xi$, where x is the end-of-period choice, and analogously to above we have $dV_t(x) = \beta^{s-t}E^{ADJ,s-t}(x)d\xi$ for all $t < s$.

At each date t, the optimal adjustment thresholds are given by value-matching conditions

$$\begin{aligned} V_t(\overline{x}_t) &= V_t(x^*_t) + \xi_t \\ V_t(\underline{x}_t) &= V_t(x^*_t) + \xi_t \end{aligned}$$

Totally differentiating these around the steady state and using $V'(x^*) = 0$, we have $d\overline{x}_t = -(dV_t(\overline{x}) - dV_t(x^*) - d\xi_t)/V'(\overline{x})$ and $d\underline{x}_t = -(dV_t(\underline{x}) - dV_t(x^*) - d\xi_t)/V'(\underline{x})$, which combined with the re-

[A-14] There are in principle firms who ended up at $x = 0$ not because they adjusted, but by chance pre-adjustment, but this is measure zero and does not affect the total number of adjusters.

sults above gives the following analogs to (23):

$$d\overline{x}_t = -\frac{\sum_{s\geq t}\beta^{s-t}\left((E^{Z,s-t}(\overline{x})-E^{Z,s-t}(x^*))dZ_s+(E^{ADJ,s-t}(\overline{x})-E^{ADJ,s-t}(x^*))d\xi_s\right)}{\sum_{s\geq t}\beta^{s-t}E^{F,s-t}(\overline{x})} \tag{A.61}$$

$$d\underline{x}_t = -\frac{\sum_{s\geq t}\beta^{s-t}\left((E^{Z,s-t}(\underline{x})-E^{Z,s-t}(x^*))dZ_s+(E^{ADJ,s-t}(\underline{x})-E^{ADJ,s-t}(x^*))d\xi_s\right)}{\sum_{s\geq t}\beta^{s-t}E^{F,s-t}(\underline{x})} \tag{A.62}$$

Similarly, the optimal reset point is given by the first-order condition $V_t'(x_t^*)=0$. Totally differentiating gives $dx_t^* = -dV_t(x^*)/V''(x^*)$, which gives the following analog to (24):

$$dx_t^* = -\frac{\sum_{s\geq t}\beta^{s-t}(E^{Z,s-t})'(x^*)dZ_s}{\sum_{s\geq t}\beta^{s-t}(E^{F,s-t})'(x^*)} \tag{A.63}$$

Substituting (A.61)–(A.63) into (A.60), we have that the Jacobian of Y with respect to Z is:

$$\begin{aligned}
&-\frac{(1-\lambda)g(\overline{x})}{\sum_{t=0}^{\infty}\beta^t E^{F,t}(\overline{x})}\begin{pmatrix} E^{Y,0}(\overline{x})-E^{Y,0}(x^*) & 0 & \cdots \\ E^{Y,1}(\overline{x})-E^{Y,1}(x^*) & E^{Y,0}(\overline{x})-E^{Y,0}(x^*) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} E^{Z,0}(\overline{x})-E^{Z,0}(x^*) & \beta(E^{Z,1}(\overline{x})-E^{Z,1}(x^*)) & \cdots \\ 0 & E^{Z,0}(\overline{x})-E^{Z,0}(x^*) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \\
&-\frac{(1-\lambda)g(\underline{x})}{\sum_{t=0}^{\infty}\beta^t E^{F,t}(\underline{x})}\begin{pmatrix} E^{Y,0}(\underline{x})-E^{Y,0}(x^*) & 0 & \cdots \\ E^{Y,1}(\underline{x})-E^{Y,1}(x^*) & E^{Y,0}(\underline{x})-E^{Y,0}(x^*) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} E^{Z,0}(\underline{x})-E^{Z,0}(x^*) & \beta(E^{Z,1}(\underline{x})-E^{Z,1}(x^*)) & \cdots \\ 0 & E^{Z,0}(\underline{x})-E^{Z,0}(x^*) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \\
&-\frac{\text{freq}}{\sum_{t=0}^{\infty}\beta^t E^{F,t\prime}(x^*)}\begin{pmatrix} E^{Y,0\prime}(x^*) & 0 & \cdots \\ E^{Y,1\prime}(x^*) & E^{Y,0\prime}(x^*) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} E^{Z,0\prime}(x^*) & \beta E^{Z,1\prime}(x^*) & \cdots \\ 0 & E^{Z,0\prime}(x^*) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}
\end{aligned} \tag{A.64}$$

The Jacobian of Y with respect to ξ is identical, but with E^Z replaced by E^{ADJ}.

Compared to equation (31) in the main text, we note that the products of lower and upper triangular matrices in (A.64) no longer have the same symmetry: in general, the sequences $E^{Y,s}(\cdot)$ and $E^{Z,s}(\cdot)$ need not be the same. Each product can be interpreted as corresponding to the pass-through matrix of a "pseudo-time-dependent" model, where agents assume a different survival function when choosing their policies than the survival function that actually governs prices. Since the symmetry between lower and upper adjustment bands is broken in the general case, there are also now three products rather than two—so that the most general form of the equivalence result away from the canonical case is that the menu cost model is equivalent to a mixture of three pseudo-time-dependent models.

E.3 Proof of proposition 4

E.3.1 Step 1: Convergence of steady state in limit $\sigma_\epsilon \to 0$

Take (A.59a) and, using the fact that the firm's profit maximization is invariant to an affine transformation, rewrite as

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t Y_t^{-\sigma} \left[\left(\frac{\zeta}{\zeta - 1} w_t \right)^{1-\zeta} Y_t \cdot \frac{F\left(x_{it} - \log W_t\right) - F(0)}{\sigma_\epsilon^2} + \xi_{it} w_t 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}} \right]$$

and then define $\hat{x}_{it} \equiv (x_{it} - \log W_{ss})/\sigma_\epsilon$ and $\hat{W}_t \equiv (\log W_t - \log W_{ss})/\sigma_\epsilon$, so that we get

$$\min_{\{\hat{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t Y_t^{-\sigma} \left[\left(\frac{\zeta}{\zeta - 1} w_t \right)^{1-\zeta} Y_t \cdot \frac{F\left(\sigma_\epsilon(\hat{x}_{it} - \hat{W}_t)\right) - F(0)}{\sigma_\epsilon^2} + \xi_{it} w_t 1_{\{\hat{x}_{it} \neq \hat{x}_{it-1} - \epsilon_{it}\}} \right]$$

Now define $\mathcal{F}(\hat{x}; \sigma_\epsilon) \equiv \frac{F(\sigma_\epsilon \hat{x}) - F(0)}{\sigma_\epsilon^2}$, so that this is just

$$\min_{\{\hat{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t Y_t^{-\sigma} \left[\left(\frac{\zeta}{\zeta - 1} w_t \right)^{1-\zeta} Y_t \cdot \mathcal{F}(\hat{x}_{it} - \hat{W}_t; \sigma_\epsilon) + \xi_{it} w_t 1_{\{\hat{x}_{it} \neq \hat{x}_{it-1} - \epsilon_{it}\}} \right] \tag{A.65}$$

Here, note that the only place that σ_ϵ enters is as a parameter to this $\mathcal{F}$ function. Further, in the limit as $\sigma_\epsilon \to 0$, it is very easy to show that $\mathcal{F}(\hat{x}; \sigma_\epsilon) \to \frac{1}{2} F''(0) \hat{x}^2$.

Explicitly, since F has a Taylor series representation around 0 and also has derivative $F'(0) = 0$, we can write

$$\lim_{\sigma_\epsilon \to 0} \mathcal{F}(\hat{x}; \sigma_\epsilon) = \lim_{\sigma_\epsilon \to 0} \frac{F\left(\sigma_\epsilon \hat{x}\right) - F(0)}{\sigma_\epsilon^2} = \lim_{\sigma_\epsilon \to 0} \frac{\frac{1}{2} F''(0) \sigma_\epsilon^2 \hat{x}^2}{\sigma_\epsilon^2} = \frac{1}{2} F''(0) \hat{x}^2$$

Note that we also get this convergence in the first derivative of $\mathcal{F}$, i.e.

$$\lim_{\sigma_\epsilon \to 0} \mathcal{F}'(\hat{x}; \sigma_\epsilon) = \lim_{\sigma_\epsilon \to 0} \frac{\sigma_\epsilon F'\left(\sigma_\epsilon \hat{x}\right) - \sigma_\epsilon F'(0)}{\sigma_\epsilon^2} = \lim_{\sigma_\epsilon \to 0} \frac{\sigma_\epsilon F''(0) \sigma_\epsilon \hat{x}}{\sigma_\epsilon^2} = F''(0) \hat{x}$$

Convergence to same steady state as quadratic objective. At the steady state, dividing both sides by $F''(0) Y^{1-\sigma} \left(\frac{\zeta}{\zeta - 1} w \right)^{1-\zeta}$, (A.65) becomes

$$\min_{\{\hat{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\mathcal{F}(\hat{x}_{it} - \hat{w}; \sigma_\epsilon)}{F''(0)} + \xi_{it} \frac{\left(\frac{\zeta}{\zeta - 1} \right)^{\zeta - 1} w^\zeta}{F''(0) Y} 1_{\{\hat{x}_{it} \neq \hat{x}_{it-1} - \epsilon_{it}\}} \right] \tag{A.66}$$

which is identical to the original steady-state optimization problem 2 with the menu cost scaled by $\frac{\left(\frac{\zeta}{\zeta - 1} \right)^{\zeta - 1} w^\zeta}{F''(0) Y}$, except that we have $\frac{\mathcal{F}(\hat{x}_{it} - \hat{w}; \sigma_\epsilon)}{F''(0)}$ rather than the quadratic objective $\frac{1}{2} \hat{x}^2$. Hence we can rewrite the recursion (A.9) for the value function in this case, and we will denote the value function by $V(\hat{x}; \sigma_\epsilon)$.

By appendix B.3, if we use the norm $\|\cdot\|$ defined in (A.17), which is a linear combination of the sup norms on V and V', $V(\hat{x};\sigma_\epsilon)$ is differentiable in σ_ϵ around $\sigma_\epsilon = 0$, with $\|dV\| \leq \frac{\|\frac{d\mathcal{F}}{F''(0)}\|}{1-\frac{1+\beta}{2}}$ where we recall that backward iteration on V is a contraction with modulus $\frac{1+\beta}{2}$ in this norm.[A-15] Analogously to appendix B.3, this differentiability implies differentiability of the policies $\underline{x}$, $\overline{x}$, and x^* in σ_ϵ, so that these all converge to their values under the quadratic objective as $\sigma \to 0$. Further, since by appendix B.2, forward iteration on beginning-of-period densities g is a contraction, and the one-period-ahead density is differentiable with respect to the policies, the steady-state density g also converges to the same as under the quadratic objective.

Aggregate price level, price dispersion, menu costs. We have verified that $\hat{x}$ approaches the same policies and distribution as $\sigma_\epsilon \to 0$. Now we also verify that all aggregate consequences are the same. In particular, we note that the aggregate price level $\frac{\zeta}{\zeta-1}\left(\int e^{(1-\zeta)\sigma_\epsilon \hat{x}_i} di\right)^{\frac{1}{1-\zeta}}$ approaches $\frac{\zeta}{\zeta-1}$ as $\sigma_\epsilon \to 0$, price dispersion $\left(\int e^{(1-\zeta)\sigma_\epsilon \hat{x}_i} di\right)^{\frac{\zeta}{1-\zeta}} \int e^{-\zeta\sigma_\epsilon \hat{x}_i} di$ approaches 1 as $\sigma_\epsilon \to 0$, and the total resources devoted to menu costs approach 0 as $\sigma_\epsilon \to 0$ (since the fraction of adjustments per period approaches a constant, but menu costs scale with σ_ϵ^2).

E.3.2 Step 2: The log-linearized system in the limit $\sigma_\epsilon \to 0$

Now, we consider the system from appendix E.1 and focus on the three equations for the firm block, $\hat{\mathbf{P}} = \mathbf{J}^{P,W}\hat{\mathbf{W}} + \mathbf{J}^{P,Y}\hat{\mathbf{Y}} + \mathbf{J}^{P,w}\hat{\mathbf{w}}$, $\hat{\mathbf{\Delta}} = \mathbf{J}^{\Delta,W}\hat{\mathbf{W}} + \mathbf{J}^{\Delta,Y}\hat{\mathbf{Y}} + \mathbf{J}^{\Delta,w}\hat{\mathbf{w}}$, $\hat{\mathbf{\Xi}} = \mathbf{J}^{\Xi,W}\hat{\mathbf{W}} + \mathbf{J}^{\Xi,Y}\hat{\mathbf{Y}} + \mathbf{J}^{\Xi,w}\hat{\mathbf{w}}$, which are the only equations directly affected by σ_ϵ. We will show that all the Jacobians in these equations converge to zero as $\sigma_\epsilon \to 0$, except for $\mathbf{J}^{P,W}$, which converges to the canonical model's pass-through matrix Ψ given the appropriately rescaled menu cost.

First, rewrite (A.65) as

$$\min_{\{\hat{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\tilde{Y}_t^{1-\sigma} \tilde{w}_t^{1-\zeta} \cdot \frac{\mathcal{F}(\hat{x}_{it} - \hat{W}_t; \sigma_\epsilon)}{F''(0)} + \tilde{Y}_t^{-\sigma} \tilde{w} \frac{\left(\frac{\zeta}{\zeta-1}\right)^{\zeta-1} w^\zeta}{F''(0)Y} \xi_{it} 1_{\{\hat{x}_{it} \neq \hat{x}_{it-1} - \epsilon_{it}\}} \right]$$

where here we denote proportional deviations from steady state by tildes, e.g. $\tilde{Y}_t \equiv Y_t/Y$. At steady state, this is the same as (A.66), which we showed is identical to the original model in steady state, with the menu cost scaled by $\frac{\left(\frac{\zeta}{\zeta-1}\right)^{\zeta-1} w^\zeta}{F''(0)Y}$. Let us redefine ξ to incorporate this scaling

[A-15] If desired, we can compute $d\mathcal{F}$ in response to $d\sigma_\epsilon$ as follows. The derivative of $\mathcal{F}(\hat{x};\sigma_\epsilon) = \frac{F(\sigma_\epsilon \hat{x}) - F(0)}{\sigma_\epsilon^2}$ with respect to σ_ϵ around $\sigma_\epsilon = 0$, using $\mathcal{F}(\hat{x};0) = \frac{1}{2}F''(0)\hat{x}^2$ for the limit, is $\lim_{\sigma_\epsilon \to 0} \frac{F(\sigma_\epsilon \hat{x}) - F(0) - \sigma_\epsilon^2 \frac{1}{2}F''(0)\hat{x}^2}{\sigma_\epsilon^3}$, which applying L'Hopital's rule once equals $\lim_{\sigma_\epsilon \to 0} \frac{\hat{x}F'(\sigma_\epsilon \hat{x}) - \sigma_\epsilon F''(0)\hat{x}^2}{3\sigma_\epsilon^2}$, applying it again equals $\lim_{\sigma_\epsilon \to 0} \frac{\hat{x}^2(F''(\sigma_\epsilon \hat{x}) - F''(0))}{6\sigma_\epsilon}$, and a third time equals $\lim_{\sigma_\epsilon \to 0} \frac{\hat{x}^3 F'''(\sigma_\epsilon \hat{x})}{6} = \frac{\hat{x}^3 F'''(0)}{6}$, so that $\left.\frac{d\mathcal{F}(\hat{x};\sigma_\epsilon)}{d\sigma_\epsilon}\right|_{\sigma_\epsilon = 0} = \frac{\hat{x}^3 F'''(0)}{6}$.

(i.e. to be $\frac{\left(\frac{\zeta}{\zeta-1}\right)^{\zeta-1}w^{\zeta}}{F''(0)Y}$ times the original ξ), so that the above becomes just

$$\min_{\{\hat{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\tilde{Y}_t^{1-\sigma} \tilde{w}_t^{1-\zeta} \cdot \frac{\mathcal{F}(\hat{x}_{it} - \hat{W}_t; \sigma_\epsilon)}{F''(0)} + \tilde{Y}_t^{-\sigma} \tilde{w}_t \xi_{it} 1_{\{\hat{x}_{it} \neq \hat{x}_{it-1} - \epsilon_{it}\}} \right] \tag{A.67}$$

For any given σ_ϵ, the first-order system in aggregates includes sequence-space Jacobians for three outcomes (the log aggregate price level $\log P_t$, price dispersion Δ_t, and menu costs incurred Ξ_t) and three shocks (log nominal wages $\sigma_\epsilon \hat{W}_t$, output $\tilde{Y}_t$, and wages $\tilde{w}_t$), all of which are given by (A.64).

We will argue, however, than in the $\sigma_\epsilon \to 0$ limit, only one of these sequence-space Jacobians remains nonzero: that for the log aggregate price level with respect to log wages. First, any change in price dispersion $\left(\int e^{(1-\zeta)\sigma_\epsilon \hat{x}_i} di\right)^{\frac{\zeta}{1-\zeta}} \int e^{-\zeta \sigma_\epsilon \hat{x}_i} di \approx 1 + \sigma_\epsilon \zeta \int \hat{x}_i di - \sigma_\epsilon \zeta \int \hat{x}_i di = 1$ is zero to first order in σ_ϵ, and menu costs paid, which scale with σ_ϵ^2, are trivially zero to first order in σ_ϵ. The $E^{Y,s}(\cdot)$ for these two outputs in (A.64) therefore scales with σ_ϵ^2 in the limit. For the log aggregate price level, on the other hand, we have $\log(\frac{\zeta}{\zeta-1}) + \frac{1}{1-\zeta} \log\left(\int e^{(1-\zeta)\sigma_\epsilon \hat{x}_i} di\right) \approx \log(\frac{\zeta}{\zeta-1}) + \sigma_\epsilon \int \hat{x}_i di$, with a nonzero term that is first order in σ_ϵ, corresponding to simple linear aggregation of price gaps. The $E^{\log P,s}(\cdot)$ for this output in (A.64) converges to σ_ϵ times the standard $E^s(\cdot)$ in the limit.

Meanwhile, for shocks to $\tilde{Y}_t$ and $\tilde{w}_t$, the $E^{Z,s}(\cdot)$ in (A.64) converge to some finite functions as $\sigma_\epsilon \to 0$.[A-16] For these, (A.64) is zero for all outcome variables, whose $E^{Y,s}(\cdot)$ went to zero as $\sigma_\epsilon \to 0$. For the shock to log nominal wages $\sigma_\epsilon \hat{W}_t$, on the other hand, we have $E^{\sigma_\epsilon \hat{W}_t,s}(\cdot) = \sigma_\epsilon^{-1} E^{\hat{W}_t,s}(\cdot)$ scaling with σ_ϵ^{-1} in the limit. For this shock, (A.64) goes to zero in the limit when the outcomes are price dispersion or menu costs (whose $E^{Y,s}(\cdot)$ scaled with σ_ϵ^2), but to a finite nonzero value when the outcome is log prices (whose $E^{Y,s}(\cdot)$ scaled with σ_ϵ, which is cancelled out by the σ_ϵ^{-1}).

We conclude that, indeed, only one sequence-space Jacobian is nonzero in the limit $\sigma_\epsilon \to 0$: that of log aggregate prices with respect to log nominal wages. Further, canceling the σ_ϵ and σ_ϵ^{-1} factors, the outcome variable is simple linear aggregation of price gaps and has a expectation function of $E^s(\cdot)$; meanwhile, the shock perturbs the loss function $\frac{\mathcal{F}(\hat{x}_{it} - \hat{W}_t; \sigma_\epsilon)}{F''(0)}$, which we have showed converges both in levels and first derivative to $\frac{1}{2}(\hat{x} - \hat{W})^2$—a quadratic function whose derivative with respect to $\hat{W}$ is unitary and also leads to an expectation function of $E^s(\cdot)$. At this point, the Jacobian (A.64) in the limit becomes identical (using symmetry and rearranging) to pass-through matrix of log marginal costs to log prices in the canonical model, as desired.

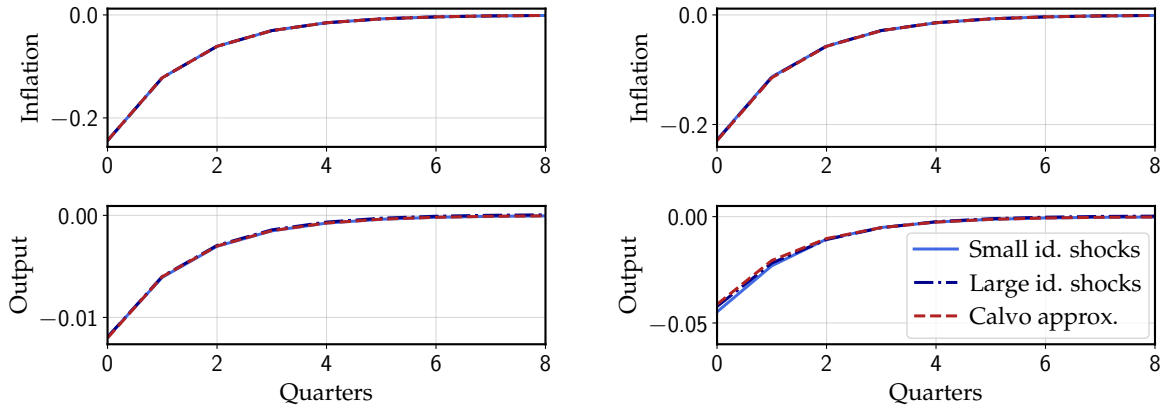
E.3.3 Step 3: Simplifying the system in the limit

Given the results in the previous subsection, the log-linearized system can be written as

$$\hat{\mathbf{P}} = \Psi \left(\hat{\mathbf{w}} + \hat{\mathbf{P}} \right)$$

$$\hat{\mathbf{w}} = (\sigma + \varphi) \hat{\mathbf{Y}}$$

[A-16] It turns out that these are zero due to the symmetry of the solution as $\sigma_\epsilon \to 0$, but we do not need this for our result.



(a) Golosov-Lucas. (b) Nakamura-Steinsson.

Figure E.1: Impulse responses with large idiosyncratic shocks

Note: both pricing models are calibrated to match the same moments as in section 2.5.

$$\hat{\boldsymbol{\pi}} = (\mathbf{I} - \mathbf{L})\,\hat{\mathbf{P}}$$

$$\hat{i}_t = \phi \hat{\pi}_t + \hat{v}_t$$

$$\hat{Y}_t = \hat{Y}_{t+1} - \frac{1}{\sigma}(\hat{i}_t - \hat{\pi}_{t+1})$$

where we note that with $\sigma_\epsilon \to 0$, the steady state price dispersion is $\Delta_{ss} = 1$ and steady state menu cost labor demand is $\Xi_{ss} = 0$. We can solve out the first equation for the price level in terms of the real wage

$$\hat{\mathbf{P}} = \Psi\,(\mathbf{I} - \Psi)^{-1}\,\hat{\mathbf{w}}$$

Combining this with the expressions for the real wage and inflation, we find

$$\hat{\boldsymbol{\pi}} = (\varphi + \sigma)\,(\mathbf{I} - \mathbf{L})\,\Psi\,(\mathbf{I} - \Psi)^{-1}\,\hat{\mathbf{Y}} = (\varphi + \sigma)\,\mathbf{K} \cdot \hat{\mathbf{Y}}$$

where $\mathbf{K}$ is exactly the generalized Phillips curve of the canonical model from section 2. Together with the Taylor rule and the Euler equation, this is the three equation system in proposition 4.

E.4 Impulse responses with large vs small σ_ϵ

Figure E.1 evaluates how well the $\sigma_\epsilon \to 0$ limit described in Proposition 4 approximates the original nonlinear model with $\sigma_\epsilon > 0$. The dashed-dark blue line labelled "large idiosyncratic shocks" shows the linear impulse response in the model with large idiosyncratic shocks, whose solution is described in section E.1. The solid-light blue line labelled "small idiosyncratic shocks" shows the approximation to that impulse response in the $\sigma_\epsilon \to 0$ limit described in proposition 4. The dashed red line shows the effects in the best-fitting Calvo model to this latter model. We calibrate the model with large idiosyncratic shocks to hit the same targets as in our calibration from section 2.5. The effects from non-zero price dispersion and aggregate menu costs, as well as the

nonlinearities in the objective function and price aggregation, are not important quantitatively.

E.5 Proof of proposition 5

In the model with strategic complementarities, firm i now produces gross output Q_{it} of variety i from hours N_{it} and intermediate X_{it} using the production function

$$Q_{it} = \frac{A_{it}}{\chi^{\chi}(1-\chi)^{1-\chi}} N_{it}^{\chi} X_{it}^{1-\chi}$$

where X_{it} is produced using the same aggregate as consumption, and has therefore the same price $P_t = \left(\int_0^1 (A_{it}P_{it})^{1-\zeta} di\right)^{\frac{1}{1-\zeta}}$.

Firm i's static profits at date t excluding menu costs are then

$$\Pi_{it} = \frac{P_{it}}{P_t} Q_{it} - \frac{MC_{it}}{P_t} \cdot Q_{it}$$

where $MC_{it} \cdot Q_{it}$ is the nominal cost for firm i of producing gross output Q_{it} at date t, with the marginal (and unit) cost MC_{it} given by

$$MC_{it} = \frac{1}{A_{it}} \underbrace{W_t^{\chi} P_t^{1-\chi}}_{\equiv MC_t}$$

where MC_t is the aggregate component of marginal cost. Factor demands are given by

$$\begin{aligned} N_{it} &= \chi \frac{1}{A_{it}} \frac{MC_t}{W_t} Q_{it} \\ X_{it} &= (1-\chi) \frac{1}{A_{it}} \frac{MC_t}{P_t} Q_{it} \\ X_{jit} &= A_{jt}^{1-\zeta} \left(\frac{P_{jt}}{P_t}\right)^{-\zeta} X_{it} \end{aligned}$$

where X_{jit} denotes firm i's demand for firm j's input at time t. Aggregating across intermediate and final good demand, total demand for firm i's output is given by

$$Q_{it} = A_{it}^{1-\zeta} \left(\frac{P_{it}}{P_t}\right)^{-\zeta} \left(\underbrace{C_t + \int X_{jt} di}_{\equiv Q_t} \right)$$

where Q_t is total gross output.

Hence, the static profits of firm i, excluding menu costs, are given by

$$\Pi_{it} = \left(\frac{P_{it}}{P_t} - \frac{1}{A_{it}} \frac{MC_t}{P_t}\right) \cdot Q_{it} = \left(\frac{P_{it}}{P_t} - \frac{1}{A_{it}} \frac{MC_t}{P_t}\right) \cdot A_{it}^{1-\zeta} \left(\frac{P_{it}}{P_t}\right)^{-\zeta} Q_t$$

and its statically optimal price is

$$\overline{P_{it}} = \frac{\zeta}{\zeta - 1}\frac{MC_t}{A_{it}} \equiv P_{it}^* \cdot MC_t$$

As in section 5.1, we can rewrite Π_{it} using $\overline{P_{it}}$ as

$$\begin{aligned}
\Pi_{it} &= \left(\frac{\zeta}{\zeta-1}\frac{MC_t}{P_t}\right)^{1-\zeta} Q_t \cdot \left(\left(\frac{P_{it}}{\overline{P_{it}}}\right)^{1-\zeta} - \frac{\zeta-1}{\zeta}\left(\frac{P_{it}}{\overline{P_{it}}}\right)^{-\zeta}\right) \\
&= \left(\frac{\zeta}{\zeta-1}\frac{MC_t}{P_t}\right)^{1-\zeta} Q_t \cdot F\left(\log\left(\frac{P_{it}}{\overline{P_{it}}}\right)\right) \\
&= \left(\frac{\zeta}{\zeta-1}\frac{MC_t}{P_t}\right)^{1-\zeta} Q_t \cdot F\left(x_{it} - \log MC_t\right)
\end{aligned}$$

where we have again defined the idiosyncratic price gap as

$$x_{it} \equiv \log P_{it} - \log P_{it}^* = \log P_{it} - \log \overline{P_{it}} + \log MC_t$$

Assuming that the menu cost is still stated in units of labor, and continuing to write $Y_t = C_t$ for GDP, the complete dynamic problem of the firm is therefore now

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t C_t^{-\sigma} \left[\left(\frac{\zeta}{\zeta-1}\frac{MC_t}{P_t}\right)^{1-\zeta} Q_t \cdot F\left(x_{it} - \log MC_t\right) + \sigma_\epsilon^2 \xi_{it} \frac{W_t}{P_t} 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}}\right]$$

with the aggregate amount of labor required for menu costs still given by (51), and labor market clearing now given by

$$N_t = \chi \frac{MC_t}{W_t}\Delta_t Q_t + \Xi_t \tag{A.68}$$

with the same equations for Ξ_t and Δ_t as in section 5.1.

Equilibrium is characterized by the same unknown sequences $\{w_t, W_t, P_t, Y_t, N_t, \Delta_t, \Xi_t, i_t, \pi_t\}$ as before, plus the 4 unknown sequences $\{Q_t, MC_t, mc_t, X_t\}$. These unknown some the same 9 equations (A.59b)–(A.59g), except that (A.59b) is replaced by a new set of functions

$$P_t = \mathcal{P}_t\left(\{MC_s, mc_s, Y_s, Q_s\}\right), \quad \Delta_t = \mathcal{D}_t\left(\{MC_s, mc_s, Y_s, Q_s\}\right), \quad \Xi_t = \mathcal{X}_t\left(\{MC_s, mc_s, Y_s, Q_s\}\right)$$

and that (A.59c) is replaced by (A.68), as well as the 4 additional equations:

$$\begin{aligned} MC_t &= W_t^{\chi} P_t^{1-\chi} \\ Q_t &= Y_t + X_t \\ X_t &= (1-\chi)\,\frac{MC_t}{P_t}\Delta_t Q_t \\ mc_t &= \frac{MC_t}{P_t} \end{aligned}$$

The proof proceeds as in section E.3. In the limit with $\sigma_\epsilon \to 0$, these new equations log-linearize as

$$\begin{aligned} \hat{mc}_t &= \hat{MC}_t - \hat{P}_t = \chi\left(\hat{W}_t - \hat{P}_t\right) = \chi\hat{w}_t \\ \hat{Q}_t &= \chi\hat{Y}_t + (1-\chi)\,\hat{X}_t \\ \hat{X}_t &= \hat{mc}_t + \hat{Q}_t \\ \hat{N}_t &= \hat{mc}_t - \widehat{w_t} + \hat{Q}_t \\ \varphi\hat{N}_t &= \hat{w}_t - \sigma\hat{Y}_t \end{aligned}$$

It can be verified that these equations simplify to

$$\begin{aligned} \hat{N}_t &= \hat{Y}_t \\ \hat{Q}_t &= \hat{Y}_t + (1-\chi)\,\hat{w}_t \\ \hat{X}_t &= \hat{Y}_t + \widehat{w_t} \\ \hat{w}_t &= (\sigma+\varphi)\,\hat{Y}_t \qquad \text{(A.69)} \\ \hat{mc}_t &= \chi\hat{w}_t \qquad \text{(A.70)} \end{aligned}$$

Hence, the relationship between real wages and output is (A.69) is the same as before, but movement in real marginal cost is scaled down by χ per (A.70). The equations characterizing equilibrium are therefore now

$$\hat{\mathbf{P}} = \Psi\left(\hat{\mathbf{mc}} + \hat{\mathbf{P}}\right)$$

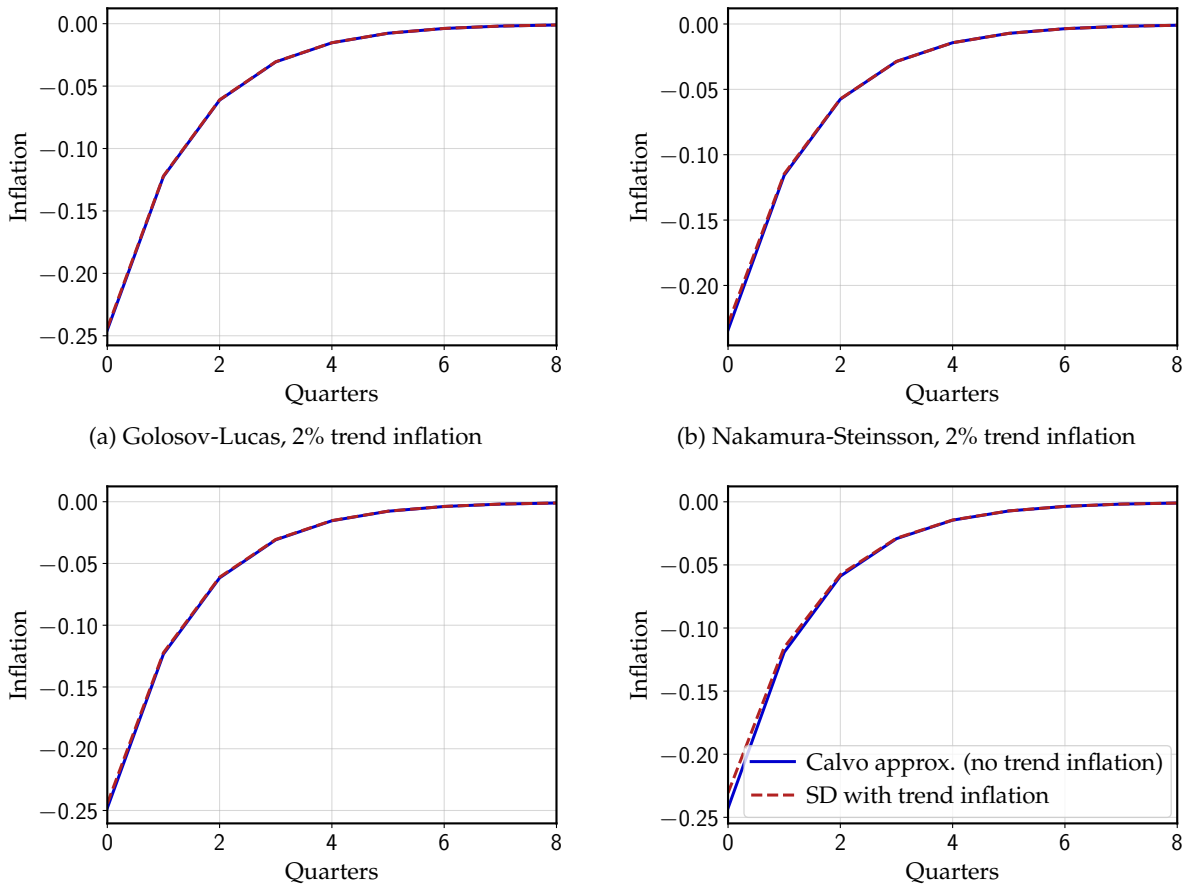
$$\hat{\mathbf{mc}} = \chi(\sigma+\varphi)\hat{\mathbf{Y}}$$

$$\hat{\boldsymbol{\pi}} = (\mathbf{I}-\mathbf{L})\,\hat{\mathbf{P}}$$

$$\hat{i}_t = \phi\hat{\pi}_t + \hat{v}_t$$

$$\hat{Y}_t = \hat{Y}_{t+1} - \frac{1}{\sigma}(\hat{i}_t - \hat{\pi}_{t+1})$$

This delivers proposition 5.



(a) Golosov-Lucas, 2% trend inflation

(b) Nakamura-Steinsson, 2% trend inflation

(c) Golosov-Lucas, 5% trend inflation

(d) Nakamura-Steinsson, 5% trend inflation

Figure E.2: Large idiosyncratic shocks and trend inflation

Note: all models are calibrated to match the same moments as in section 2.5, except for higher trend inflation. Inflation is in annualized terms.

E.6 Trend inflation in the general equilibrium model

Here, we extend the general equilibrium model with large idiosyncratic shocks from section E.4 to feature trend inflation. We compare the impulse responses of inflation to a monetary policy shock generated by our two state dependent models (GL and NS) to what one would obtain from a Calvo model without trend inflation designed to approximate the generalized Phillips curve of each menu cost model. Figure E.2 shows results. The inflation responses generated by the menu cost model are still indistinguishable from those of a Calvo model.[A-17]

[A-17] Interestingly, with trend inflation, we noticed a somewhat larger departure in the output responses to a monetary shock, coming from a negative effect of monetary tightening on the frequency of price changes, and hence on the aggregate demand coming from menu cost payments.

E.7 The Calvo model with trend inflation

Here we compute the pass-through matrix of a Calvo model with trend inflation. More specifically, the model we analyze here is the same one as in section 5, but with (i) infinitely large menu costs, so that firms only adjust prices when they receive a free adjustment opportunity, (ii) no idiosyncratic shocks ($\sigma = 0$) and (iii) a trend μ in the law of motion of the log nominal wage (and, more generally, all nominal variables).

We use the model from section 5 because trend inflation does not affect the pass-through matrix of the Calvo model when profits are assumed to be quadratic in the price gap x. Therefore, we need the profit function (48). Moreover, for computing the pass-through matrix, we can discard the terms that do not interact with the price gap and assume that profits are given simply by

$$F(x) \equiv e^{(1-\zeta)x} - \frac{\zeta - 1}{\zeta} e^{-\zeta x}.$$

Given a sequence of nominal wages $\{\log W_t\}_{t=0}^{\infty}$, already expressed as deviations from the long-run trend, firms that are allowed to reset their prices choose a price gap that solves

$$x_t^* = \arg\max_x \sum_{k=0}^{\infty} (\beta(1-\lambda))^k F(x - \mu t - \log W_t), \tag{A.71}$$

where λ is the adjustment probability and μ is the long-run inflation trend. The price level is still given by equation (49), which now becomes

$$\begin{aligned} \log P_t &= \log\left(\frac{\zeta}{\zeta - 1}\right) + \frac{1}{1-\zeta} \log\left(\int e^{(1-\zeta)x_{it}} di\right) \\ &= \log\left(\frac{\zeta}{\zeta - 1}\right) + \frac{1}{1-\zeta} \log\left(\sum_{k=0}^{\infty} \lambda(1-\lambda)^k \exp\left((1-\zeta)(x_{t-k}^* - \mu k)\right)\right) \end{aligned} \tag{A.72}$$

Without trend inflation, the pass-through matrix of a Calvo model is characterized by two parameters: the degree of price stickiness and the discount factor, and can be computed using equation (12). With trend inflation, we have the following result:

Proposition 9. *The pass-through matrix of the Calvo model with trend inflation is given by*

$$\Psi = \zeta \Psi_1 + (1-\zeta)\Psi_2,$$

where Ψ_1 and Ψ_2 are pass-through matrices of Calvo models without trend inflation. Both are characterized by the same adjustment probability

$$\tilde{\lambda} = 1 - (1-\lambda)e^{(\zeta-1)\mu},$$

but differ in the discount factor: Ψ_1 is associated with the discount factor $\tilde{\beta} = \beta e^{\mu}$, while Ψ_2 is associated with the original discount factor β.

Before proving this proposition, it is useful to highlight two features of this result. First, it is

straightforward to see that we obtain the standard Calvo pass-through matrix by setting $\mu = 0$, in which case $\Psi = \Psi_1 = \Psi_2$. Second, the weight of the first component is the elasticity of substitution $\zeta > 1$, so even though weights sum to one, this is not a convex combination of two standard Calvo models, as Ψ_2 is associated with a negative weight.

The proof of the proposition above is as follows. We first need to compute the steady state reset point x^*. Differentiate (A.71) to obtain the first order condition

$$\sum_{k=0}^{\infty}(\beta(1-\lambda))^t \left[e^{(1-\zeta)(x_t^* - \log W_{t+k} - k\mu)} - e^{-\zeta(x_t^* - \log W_{t+k} - \mu t)}\right] = 0. \tag{A.73}$$

In the steady state characterized by $\log W_t = 0$ for all t, this becomes

$$x^* = \log\left(\frac{1-\beta(1-\lambda)e^{(\zeta-1)\mu}}{1-\beta(1-\lambda)e^{\zeta\mu}}\right). \tag{A.74}$$

From this equation, we can see that there is a maximum level for μ, above which x^* goes to infinity. The intuition is that losses associated with negative price gaps are unbounded, while profits go to zero as $x \to +\infty$. Firms, therefore, avoid the potentially unbounded losses that occur if they remain long periods without adjusting by setting very large reset points x^* as trend inflation increases.

Now, we compute the response of the price level $\log P_t$ to past pricing decisions x_{t-s}^*, analogous to the matrix on the right hand side of (12). Using equation (A.72), we obtain

$$\frac{\partial \log P_t}{\partial x_{t-k}^*} = \lambda(1-\lambda)^k \frac{e^{(1-\zeta)(x_{t-k}^* - \mu k)}}{\sum_{k=0}^{\infty}\lambda(1-\lambda)^k e^{(1-\zeta)(x_{t-k}^* - \mu k)}}.$$

Around steady state, this becomes

$$\begin{aligned}\frac{\partial \log P_t}{\partial x_{t-k}^*} &= \left(1-(1-\lambda)e^{(\zeta-1)\mu}\right)\left((1-\lambda)e^{(\zeta-1)\mu}\right)^k \\ &= \tilde{\lambda}(1-\tilde{\lambda})^k,\end{aligned} \tag{A.75}$$

for $\tilde{\lambda} = 1-(1-\lambda)e^{(\zeta-1)\mu}$, as defined above, which is exactly what one obtains from a Calvo model with adjustment probability $\tilde{\lambda}$.

The last step is to compute the response of the reset price gap to a shock to future nominal marginal costs: $\frac{\partial x_t^*}{\partial \log W_{t+s}}$. By differentiating (A.73) with respect to $\log W_{t+s}$ around steady state, we obtain

$$\frac{\partial x_t^*}{\partial \log W_{t+s}} = \frac{(\beta(1-\lambda))^s\left[\zeta e^{\zeta\mu s} - (\zeta-1)e^{x^* + (\zeta-1)\mu s}\right]}{\sum_{k=0}^{\infty}(\beta(1-\lambda))^k\left[\zeta e^{\zeta\mu k} - (\zeta-1)e^{x^* + (\zeta-1)\mu k}\right]}.$$

This can be rewritten as

$$\frac{\partial x_t^*}{\partial \log W_{t+s}} = \frac{\left[\zeta\left((1-\tilde{\lambda})\tilde{\beta}\right)^s + (1-\zeta)e^{x^*}\left((1-\tilde{\lambda})\beta\right)^s\right]}{\sum_{k=0}^{\infty}(\beta(1-\lambda))^k\left[\zeta e^{\zeta\mu k} - (\zeta-1)e^{x^*+(\zeta-1)\mu k}\right]},$$

where $\tilde{\beta} = \beta e^{\mu}$, as defined above.

Some manipulation using the expression from steady state reset point (A.74) gives us

$$\frac{\partial x_t^*}{\partial \log W_{t+s}} = \zeta\left(1-(1-\tilde{\lambda})\tilde{\beta}\right)\left((1-\tilde{\lambda})\tilde{\beta}\right)^s + (1-\zeta)\left(1-(1-\tilde{\lambda})\beta\right)\left((1-\tilde{\lambda})\beta\right)^s. \tag{A.76}$$

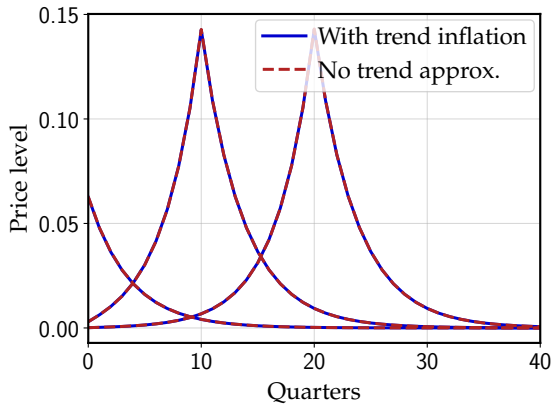
Note that each term on the right hand side corresponds to what one would obtain from a standard Calvo model, except for the adjustment in the first discount factor. The result then follows from writing equations (A.75) and (A.76) in matrix form, as in (12).

We now ask whether the resulting pass-through matrix is different from the one implied by a Calvo model without trend inflation. Similarly to section 4, we compute the pass-through matrix for the Calvo model with many levels of trend inflation and then compute the best approximating Calvo model *without* trend inflation. We calibrate $\lambda = 0.25$ and $\beta = 0.95$. We employ a relatively low value for β here to avoid invertibility issues when computing the generalized Phillips curve. Moreover, when computing the Calvo approximation without trend inflation, we allow the discount factor to vary as well.

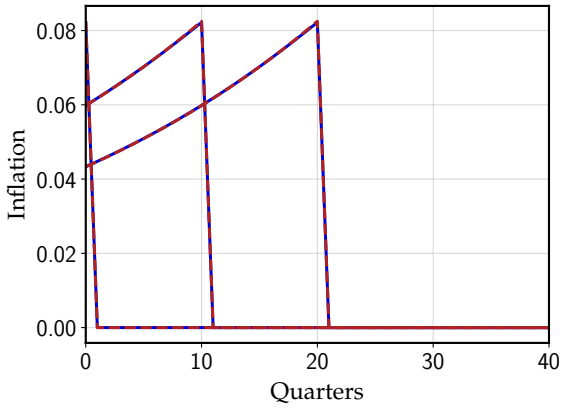
Figure E.3 shows results. For moderate inflation levels, the pass-through matrix and the generalized Phillips curve of the Calvo model with trend inflation is very similar to the one obtained *without* trend inflation, as long as we allow the discount factor to vary, as shows in the top two panels. In the bottom two panels, we show the necessary discount factor and slope of the Phillips curve required to approximate the Calvo model with trend inflation. As suggested by the results above, as trend inflation increases, we need a lower discount factor, as well as a flatter Phillips curve, to approximate the generalized Phillips curve.

Finally, we analyze how menu cost and Calvo models behave differently as we increase trend inflation. To do so, we start with a menu cost model with large idiosyncratic shocks and *without* trend inflation, as in section D.5.5, and compute its generalized Phillips curve and the corresponding Calvo approximation (without trend inflation). Then, we add 2% annualized trend inflation to both models, Calvo and menu cost, and compute the corresponding generalized Phillips curves.

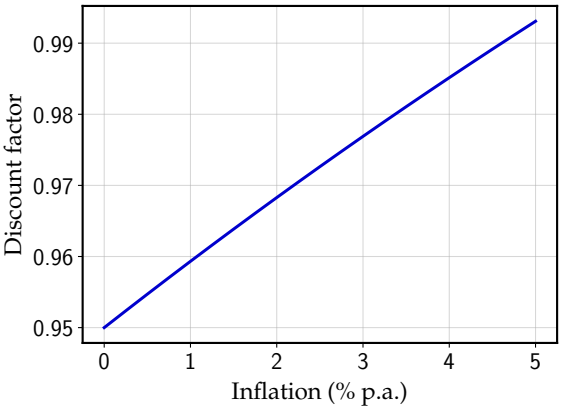
Figure E.4 shows the results. The blue line shows the generalized Phillips curve of the menu cost model with trend inflation, while the red and green lines show the Phillips curves of the Calvo models with and without trend inflation, respectively. Quite surprisingly, we find that trend inflation of this magnitude only has a small effect in the menu cost model, echoing our results in appendix D.5.1; by contrast, trend inflation has a large effect within the Calvo model. This is because the inability to endogenously adjust prices in the Calvo model makes firms susceptible to arbitrarily large losses as we increase trend inflation. We end up with the seemingly paradoxical result that the standard Calvo NK-PC describes better a menu cost model with trend inflation than



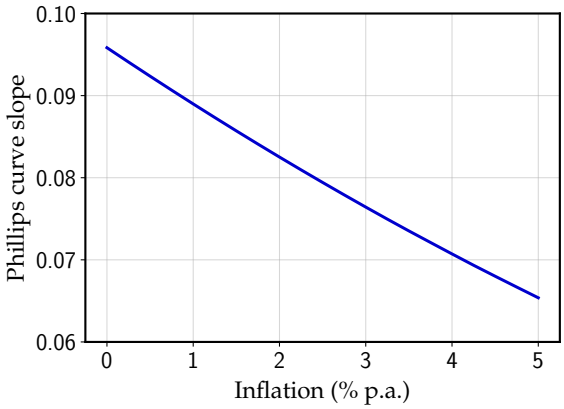
(a) Pass-through matrices for $\mu = 0.02$



(b) Generalized Phillips curves for $\mu = 0.02$



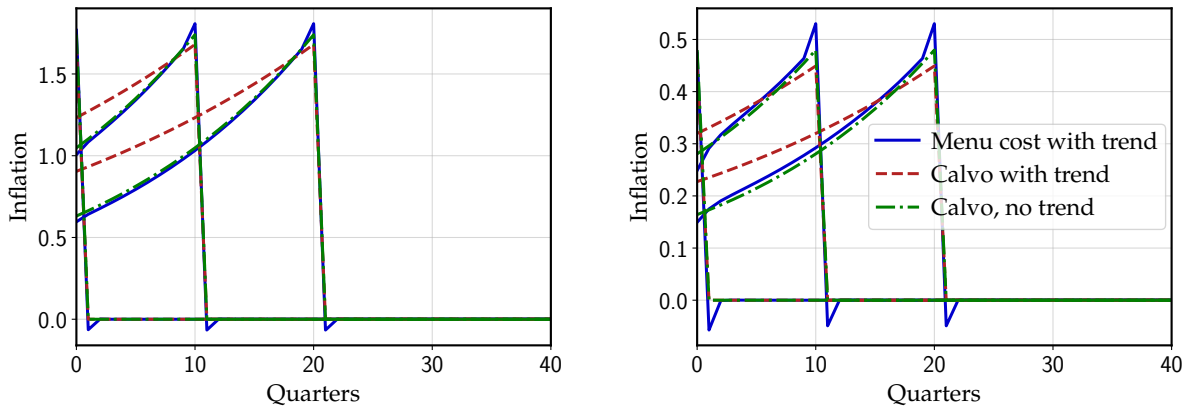
(c) Discount factor of the best-fitting Calvo model



(d) Slope of the best-fitting Calvo Phillips curve

Figure E.3: Calvo model with trend inflation approximated by a standard Calvo model

Note: the top panels show the pass-through matrices and generalized Phillips curves for a Calvo model with 2% annual trend inflation, as well as the best-fitting Calvo model without trend inflation. The bottom panels show the discount factor and the slope of the best-fitting Calvo model as a function of trend inflation μ.



(a) Golosov-Lucas (b) Nakamura-Steinsson

Figure E.4: Comparing Calvo and menu cost models with trend inflation

Note: for a given menu cost model without trend inflation, the figure shows the generalized Phillips curve of its Calvo approximation (green), as well as the generalized Phillips curves of the menu cost model (blue) and the Calvo approximation (red) after adding a 2% annual trend inflation. The menu cost models are calibrated to match the moments in 2.5. The discount factor is set to $\beta = 0.95$.

a Calvo model with trend inflation.

F Appendix to Section 6

F.1 Proof of proposition 6

We now assume that the menu cost ξ_{it} is drawn from an arbitrary distribution with differentiable c.d.f $\mathcal{H}(\cdot)$. In period t, a firm with price gap x adjusts its price with probability

$$\Lambda_t(x) = P\left(\xi_{it} \leq V_t(x) - V_t(x_t^*)\right) = \mathcal{H}\left(V_t(x) - V_t(x_t^*)\right). \tag{A.77}$$

Now expected price gaps evolve according to

$$E^{t+1}(x) = \int_{-\infty}^{\infty} \left(1 - \Lambda(x')\right) f(x' - x) E^t(x') dx'$$

and the steady state adjustment frequency is given by

$$\text{freq} = \int_{-\infty}^{\infty} \Lambda(x) g(x) dx,$$

where $g(x)$ is the steady state distribution of price gaps.

We first characterize the extensive margin price level response dP_t^e generated by a change in

adjustment probabilities $\{d\Lambda_s(x)\}_{s=0}^{\infty}$. Similarly to the first part of equation (25), we have

$$dP_t^e = -\sum_{s=0}^{t}\int_{-\infty}^{\infty}\left[d\Lambda_{t-s}(x)g(x)E^s(x)\right]dx. \tag{A.78}$$

Intuitively, a perturbation in adjustment probabilities $d\Lambda_{t-s}(x)$ generates an additional mass of price changes $d\Lambda_{t-s}(x)g(x)$, which then changes the price level at date t by $-d\Lambda_{t-s}(x)g(x)E^s(x)$. The intensive margin response dP_t^i is still given by the second part of equation (25)

The next step is to characterize the responses of optimal policies $\Lambda_t(x)$ and x_t^* to aggregate marginal cost shocks. First, notice that the reset point dynamics is still given by (30). Now differentiate (A.77) with respect to x and evaluate it at steady state to obtain

$$\Lambda'(x) = -h\left(V(0) - V(x)\right)V'(x),$$

where $h = \mathcal{H}'$. By totally differentiating (A.77), also around steady state, one gets

$$\begin{aligned} d\Lambda_t(x) &= h\left(V(0) - V(x)\right)\left(V'(0)dx_t^* + dV_t(0) - dV_t(x)\right) \\ &= -\Lambda'(x)\frac{dV_t(0) - dV_t(x)}{V'(x)}, \end{aligned}$$

where the second line uses $V'(0) = 0$. One can still obtain $V'(x) = \sum_{u=0}^{\infty}\beta^u E^u(x)$ and $dV_t(x)$ from equation 28. Using the definition $\Phi_t^e(x) = E^t(x)/\overline{x}$, we have

$$-\frac{d\Lambda_t(x)}{\Lambda'(x)} = \frac{\sum_{s=t}^{\infty}\beta^{s-t}\Phi_{t-s}^e(x)d\log MC_s}{\sum_{s=t}^{\infty}\beta^{s-t}\Phi_{t-s}^e(x)}. \tag{A.79}$$

This implies that $-d\Lambda_t(x)/\Lambda'(x)$ responds to future marginal costs according to weights $\beta^t\Phi_t^e(x)$, just like the reset point of a TD model in (10).

Now rewrite (A.78) as

$$dP_t^e = \sum_{s=0}^{t}\int_{-\infty}^{\infty}\left[\Lambda'(x)g(x)\left(\sum_{\tau=0}^{\infty}E^\tau(x)\right)\frac{\sum_{s=0}^{t}\Phi_s^e(x)\left(-\frac{d\Lambda_{t-s}(x)}{\Lambda'(x)}\right)}{\sum_{\tau=0}^{\infty}\Phi_\tau^e(x)}\right]dx.$$

This shows that dP^e responds to changes in past policies $-d\Lambda_t(x)/\Lambda'(x)$ with weights $\Phi_t^e(x)$, as the price level of a TD model responds to changes in past pricing decisions in (11). It follows from this that the extensive margin dynamics of the price level is equivalent to the sum of a continuum of TD models, one for each point x, with survival function $\Phi_t^e(x)$ and weight $\Lambda'(x)g(x)\left(\sum_{\tau=0}^{\infty}E^\tau(x)\right)$. Proposition 6 then follows from $d\hat{P}_t = dP_t^e + dP_t^i$.

F.2 Details on the section 6 algorithm and measurement

Here we provide further details on the algorithm used to construct the pass-through matrix in (58), and on the empirical implementation.

F.2.1 Algorithm

In order to evaluate (58), we need to measure four different objects:

- the frequency of adjustment freq.
- the expected price gap function $E^t(x)$, for any x, which then also determines $\Phi_t^i = E^{t\prime}(0)$ and $\Phi^e(x) = \frac{E^t(x)}{x}$.
- the generalized hazard function $\Lambda(x)$.
- the stationary distribution $g(x)$.

As an auxiliary object, we will also measure the variance of idiosyncratic shocks σ_ϵ^2. The data we assume a researcher has access to is (a) the empirical frequency of price adjustments; (b) the distribution of price changes $\{\Delta p_i\}$ for all products i whose price changes.

Obtaining σ_ϵ^2 and $\Lambda(x)$. To obtain the generalized hazard function $\Lambda(x)$, we use the following algorithm:

- Guess σ_ϵ and $\Lambda(x)$.[A-18]
- Given σ_ϵ and $\Lambda(x)$, compute the steady state density of price gaps $g(x)$.
- Compare the observed density of price changes Δp with the theoretical one, $\Lambda(-\Delta p)g(-\Delta p)$, and update guesses for σ_ϵ and $\Lambda(x)$ to find the least squares fit.

In practice, to operationalize the algorithm, we choose a parametric functional form for $\Lambda(x)$. We have found that the following functional form works well:

$$\log\left(\frac{\Lambda(x)}{1-\Lambda(x)}\right) = p(x) - s\phi_\sigma(x) \tag{A.80}$$

In the expression above, $p(x) = p_0 + p_2 x^2$ is an even polynomial of degree 2, $\phi_\sigma(\cdot)$ is the p.d.f. of a normal distribution with standard deviation σ, and s is a scaling factor. The scaled normal p.d.f. generates the drop in adjustment hazards close to $x = 0$, visible in figure 13, necessary for matching the price-change distribution.

[A-18] One alternative is to use the identity $\text{Var}(\Delta p) = \sigma_\epsilon^2$, which states that the steady-state variance of all price changes (including zeros) equals σ_ϵ^2, to directly obtain σ_ϵ^2 from the data. This identity can be obtained in the model and is similar to derivations in Alvarez et al. (2016). We opt instead to solve jointly for Λ and σ_ϵ^2 to achieve a best fit.

To compute the steady-state density of price gaps $g(x)$, we iterate on the condition for steady-state g

$$g(x) = f(x)\left(\int_{-\infty}^{\infty}\Lambda(x')g(x')dx'\right) + \int_{-\infty}^{\infty} f(x-x')(1-\Lambda(x'))g(x')dx' \tag{A.81}$$

where f is the density of innovations ϵ, normal with mean 0 and variance σ_ϵ^2. In (A.81), the right term is the contribution to density of non-resetters, and the left term is the contribution of resetters, with the term in parentheses being freq.

In practice, we choose an interval $[-\bar{x}, \bar{x}]$ on which to calculate g that is wide enough that Λ is nearly 1 and density is effectively zero outside of it. We then represent g as a cubic spline through a large number of equispaced nodes (~200) on the interval. Our initial guess for g can be an arbitrary symmetric function (say, uniform on $[-\bar{x}, \bar{x}]$). We then plug it in to the right of (A.81) and evaluate at all spline nodes x, computing the integral on the right using Gauss-Hermite quadrature with relatively few nodes (~7) and the integral on the left using Gauss-Legendre quadrature with relatively many nodes (~50) on the interval $[0, \bar{x}]$ and then doubling, taking advantage of symmetry.[A-19] We re-interpolate the results to obtain a new cubic spline g, normalize by the integral of g (calculated directly from the cubic spline representation) to ensure that numerical error does not cause any small leakage of probability mass, and repeat the process until densities at successive iterations differ by a maximum that is below some convergence criterion (~10^{-8}).

Given the resulting $g(x)$, the density of price changes $\Lambda(\Delta p)g(\Delta p)$ follows immediately. We compare this density to the empirical density of price changes (scaled by frequency) and then repeat the process, solving for the five parameters (σ_ϵ^2, and four parameters for $\Lambda(x)$) that achieve the least-squares fit to the empirical density using the Levenberg-Marquardt algorithm.

The empirical density could come in multiple forms, but in our case it comes in the form of a histogram with equal bin sizes. For our comparison between model and empirical density, we simply compare the model's price change density at the center of each bin with the appropriately scaled histogram.

Obtaining $E^t(x)$. Once $\Lambda(x)$ is determined, $E^t(x)$ can be obtained through the iteration

$$E^t(x) = \int_{-\infty}^{\infty} f(x'-x)\left(1-\Lambda(x')\right)E^{t-1}(x')dx' \tag{A.82}$$

which directly generalizes (19) for the case of a generalized hazard function $\Lambda(x)$. Analogously to (A.81), we represent each $E^t(x)$ by a cubic spline over many nodes on $[-\bar{x}, \bar{x}]$, and calculate the integral in (A.82) using Gauss-Hermite quadrature.

[A-19] A large number of nodes is needed for the latter because the function $\Lambda(x')g(x')$ is not well-approximated on the full inteval by a low-degree polynomial.

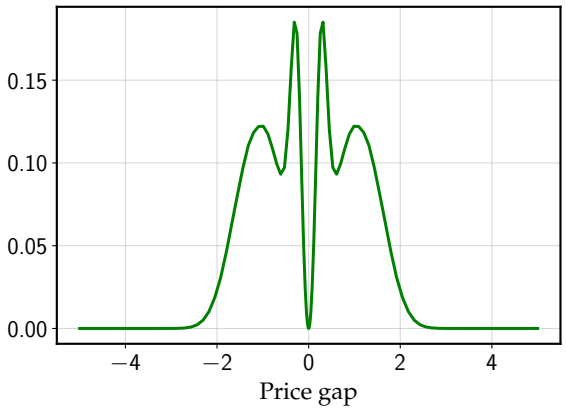


Figure F.1: Extensive margin weights in the application to Israel

Note. This is the distribution over the extensive margin weights $\Lambda'(x)g(x) \cdot \left(\sum_{t=0}^{\infty} E^t(x)\right)$ in the generalized exact equivalence result (58).

F.2.2 Implementation on data

Here we briefly describe the dataset used in section 6, as well as the numerical implementation. We use data from Bonomo et al. (2022). A law enacted in 2014 requires large food retailers in Israel to post online information on the prices of all their products on a daily basis, which the Bank of Israel then collects. The only empirical object we use is the price-change distribution in figure 13, which is computed using data on the top 5% stores in terms of number of observations, totaling 506.1 million daily observations. The size of each price change is standardized by the within-store standard deviation of price changes in order to filter out store heterogeneity. This does not affect the pass-through matrix, which is invariant to rescaling of all price gaps.

Figure F.1 shows the extensive margin weights $\Lambda'(x)g(x) \cdot \left(\sum_{t=0}^{\infty} E^t(x)\right)$ as function of the price gap x. The weights vanish close to the reset point at $x = 0$ as well as for price gaps x further away from the reset point. In both of these regions, the generalized adjustment hazard has a zero derivative, $\Lambda'(x) = 0$. Otherwise, the weights are positive.