New Pricing Models, Same Old Phillips Curves?

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

April 2022

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 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, 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 by Mark Aguiar, Fernando Alvarez, Marios Angeletos, Luigi Bocola, Marco Bonomo, Ricardo Caballero, Luca Dedola, Yoon Jo, Pete Klenow, Oleksiy Kryvtsov, John Leahy, 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.

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:

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

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 all the first-order aggregate implications of the Calvo price-setting model. 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 availability of an equivalent to the New Keynesian Phillips curve.[4]

[1] Another widely used example of a time-dependent model is Taylor (1979). 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, including Barro (1972) and Sheshinski and Weiss (1977).

[4] Instead, to study aggregate implications of menu cost models, the literature has had to resort to stark general equilibrium assumptions, such as the combination of one-time permanent money shocks and a utility function for the

In this paper, we fill this gap. We extend the notion of an aggregate 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 equivalence* between SD and TD. Second, for a wide range of common parameterizations, this Phillips curve is numerically almost identical to the Calvo Phillips curve (NK-PC), for some κ. This *numerical 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 general price-setting models. For any such model, we define two new objects, both in the form of Jacobian matrices relating impulse responses.[5] The first is the *pass-through matrix* Ψ, characterizing the first-order mapping from arbitrary shocks to nominal marginal cost $\widehat{\mathbf{MC}} = \left(\widehat{MC}_0, \widehat{MC}_1, \ldots\right)$ to the entire impulse response of aggregate prices $\hat{\mathbf{P}} = \left(\hat{P}_0, \hat{P}_1, \ldots\right)$. The second matrix is the *generalized Phillips curve* $\mathbf{K}$, characterizing the first-order mapping from arbitrary shocks to real marginal cost $\widehat{\mathbf{mc}} = \left(\widehat{mc}_0, \widehat{mc}_1, \ldots\right)$ to the entire impulse response of inflation $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots)$. The matrices Ψ and $\mathbf{K}$ are themselves related by a simple one-to-one mapping. We characterize both explicitly for TD models, and then proceed to analyze them for menu cost models.

Our first main result characterizes Ψ and $\mathbf{K}$ 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, permanent idiosyncratic productivity shocks, and random opportunities for free adjustments. We prove that the pass-through matrix Ψ of the canonical model is a convex combination of the pass-through matrices of two TD models. Hence, the first-order 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 permanent shock to marginal cost as the sum of an extensive and extensive margin component. Our result shows how their structural decomposition extends to the entire impulse response of prices and to arbitrary 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.

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

representative agent that features linear disutility of labor, and log utilities over consumption and money.

[5] These matrices are known as "sequence-space Jacobians" (Auclert, Bardóczy, Rognlie and Straub 2021).

shock to their ideal price, or a shock drawn from a uniform distribution with wide support. Given this assumption, 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 *numerical equivalence* to a Calvo model. 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. This 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.

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 New Keynesian Phillips curve (NK-PC) 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.[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 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 shock to costs.[7] The two results are highly complementary: the Alvarez, Le Bihan and Lippi (2016) sufficient statistic allows to compute 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, all impulse responses are numerically close between the two models.

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 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 (NK-PC) with the generalized Phillips

[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.

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 (NK-PC) 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 to use this extended result to directly compute the pass-through matrix and generalized Phillips curve from the empirical distribution of price changes alone, without any need for model simulation. Applying this technique to Israeli supermarket data from Bonomo, Carvalho, Kryvtsov, Ribon and Rigato (2020), we again find that the resulting generalized Phillips curve is 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 $\mathbf{K}$: using our exact equivalence result (e.g. from impulse responses to permanent shocks alone), using our approximate equivalence result (finding the κ that provides the best fit), or using the price change distribution directly. Second, for the literature developing DSGE models, we provide a new rationalization of the Calvo Phillips curve based on menu costs without the need to resort to special assumptions.

Finally, for the literature developing price-setting models that can match both micro and macro data, we provide both optimism and caution. Optimism, because menu cost models can be represented using generalized Phillips curves that can be taken to the macro data. Caution, because these Phillips curves are so close to the Calvo model that they suffers from the same deficiencies, such as a lack of internal inflation persistence (Fuhrer and Moore 1995) and extreme forward-looking behavior. One has to look beyond menu cost models alone[9] to resolve these puzzles: for instance, to multi-sector models with complex input-output linkages (Rubbo 2020, La'O and Tahbaz-Salehi 2020), or to deviations from full-information rational expectations (Mankiw and Reis 2002, Woodford 2003a, Sims 2003, Nimark 2008, Maćkowiak and Wiederholt 2015, Gabaix 2020, Angeletos and Huo 2021, Afrouzi and Yang 2021).

Our results establish a first-order connection between menu cost and TD models,[10] and so should not be taken to imply equivalence between menu costs and Calvo beyond first order in aggregates. For instance, any dependence on the aggregate state implied by SD models (as stressed, for instance, by Caballero and Engel 1992 and Vavra 2014) is not captured by this equivalence, and the welfare implications of both models may be quite different. We leave an exploration of

[9]In particular, our results shows that the intuition according to which the increasing hazards in menu cost models may generate inflation persistence because they do so in TD models (Sheedy 2010) is incorrect. This is because it is the virtual hazards, rather than the observed hazards, that matter for inflation persistence.

[10]This connection transcends price-setting applications. For instance, it also applies to investment with fixed costs.

the question of higher-order properties of menu cost models to future research.

In parallel and independent work, Alvarez, Lippi and Souganidis (2022b) also study the pass-through matrix Ψ of the canonical menu cost, focusing on the case of Brownian shocks in continuous time. The main goal of their paper is to characterize the impulse response to a permanent nominal cost shock under strategic complementarities. This is closely related to our study of the generalized Phillips curve $\mathbf{K}$, since a shock to real cost acts like a strategic complementarity from the point of view of nominal price-setting. In fact, we show that adding a standard type of real rigidity to our model simply amounts to scaling down $\mathbf{K}$ by a constant. Alvarez et al. (2022b) primarily use their expression for Ψ in terms of model primitives to show that it is self-adjoint, while we use ours to study the equivalence to a time-dependent model and obtain sufficient statistics for Ψ. Together, our papers show the importance of the pass-through matrix for the general equilibrium solution of menu cost models, either when shocks to nominal costs are not permanent or when there are strategic complementarities.

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 pass-through matrix and 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.[11]

[11]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 $\log P_t = \sum_{j=0}^{\infty} \widehat{P}_j \epsilon_{t-j}$. 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) and Fernández-Villaverde, Rubio-Ramírez and Schorfheide (2016). 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).

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, and random opportunities for free adjustments. 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. 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 proportional to $\frac{1}{2}\left(p_{it} - p_{it}^* - \log MC_t\right)^2$, 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 price setting 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 that definition, firm i solves the following pricing problem:

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

where the menu cost ξ_{it} has to be paid whenever x_{it} is chosen to differ from $x_{it-1} - \epsilon_{it} = p_{it-1} - p_{it}^*$, or in other words, whenever $p_{it} \neq p_{it-1}$.

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 a first-order approximation of a microfounded price-setting problem with menu costs.

As we show formally in appendix B, the solution to problem (2) has a well-known "sS" pattern, with the optimal policy taking the form

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

Following the literature, we refer to $\underline{x}_t$ and $\overline{x}_t$ as the *lower* and *upper adjustment barriers*, and x_t^* as the *reset point*. The triplet $(\underline{x}_t, \overline{x}_t, x_t^*)$ constitutes the *policies* of the menu cost model. In general, these policies vary over time when the sequence of nominal marginal cost $\log MC_t$ does. In a steady state $\log MC_t$ is constant, and we can normalize it to $\log MC_t = 0$. 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 $x_{it-1} - \epsilon_{it}$ *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. Price resets come from firms whose prices have fallen beyond the adjustment barriers and free resets inside the adjustment band, so $\text{freq} = \int_{-\infty}^{\underline{x}} g(x)dx + \int_{\overline{x}}^{\infty} g(x)dx + \lambda \int_{\underline{x}}^{\overline{x}} g(x)dx$.

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). In particular, 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. Mechanically, $\Phi_0 = 1$, and $\Phi_s \in [0,1]$ is weakly decreasing in s.[12]

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

$$x_t^* \equiv \arg\min_x \mathbb{E}_0 \sum_{s=0}^{\infty} \beta^s \Phi_s \frac{1}{2} \left(x - \sum_{r=1}^{s} \epsilon_{it+r} - \log MC_{t+s} \right)^2 \tag{4}$$

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 (4) 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

[12] For proposition 3 we will impose the regularity condition that survival eventually declines exponentially, i.e. that $\Phi_s < C/v^s$ for some constant C and $v > 1$.

expressed as

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

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 then 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 $a_s = (1-\lambda_s)\, a_{s-1}$, which combined with the definition of λ_s delivers $a_s \propto \Phi_s$. Since $\sum_{s=0}^{\infty} a_s = 1$, we find the share of prices with age s to be

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

2.3 Aggregate dynamics: pass-through matrix

The SD and TD models defined so far have in common that, at the aggregate level, they translate a sequence of nominal marginal cost $\{MC_t\}$ 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

$$P_t = \mathcal{P}_t\left(\{MC_s\}\right) \tag{5}$$

In this paper, we almost exclusively focus on small (first-order) shocks to marginal cost. Log-linearizing (5), and denoting log deviations with a "hat", we have

$$\hat{P}_t = \sum_{s=0}^{\infty} \frac{\partial \log \mathcal{P}_t}{\partial \log MC_s} \widehat{MC}_s \tag{6}$$

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 $\Psi_{t,s} \equiv \frac{\partial \log \mathcal{P}_t}{\partial \log MC_s}$. Stacking (6) across t into a vector-valued equation, we obtain

$$\hat{\mathbf{P}} = \Psi \cdot \widehat{\mathbf{MC}} \tag{7}$$

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, (7) can be used to evaluate the impulse response of the price level $\hat{\mathbf{P}}$ with respect to an arbitrary (first order) nominal marginal cost shock $\widehat{\mathbf{MC}}$. For example, 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 corresponding to (4),

$$x_t^* = \frac{\sum_{s\geq 0} \beta^s \Phi_s \widehat{MC}_{t+s}}{\sum_{s\geq 0} \beta^s \Phi_s} \tag{8}$$

Equation (8) 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 beta-discounted version of the survival function Φ_s. We refer to (8) as the *policy equation*.

From (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,

$$\hat{P}_t = \sum_{s\geq 0} a_s x_{t-s}^* = \frac{\sum_{s\geq 0} \Phi_s x_{t-s}^*}{\sum_{s\geq 0} \Phi_s} \tag{9}$$

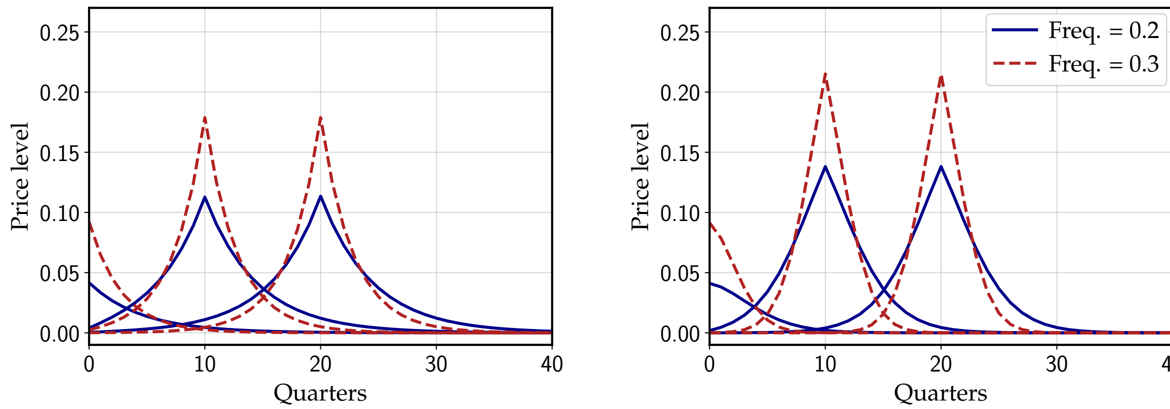
We refer to (9) as the *law of motion* of the price level. Notice that the weights in the policy equation (8) are exactly the beta-discounted versions of the weights that appear in the law of motion (9). This is a key property of TD models to which we will return.

Combining the policy equation (8) and the law of motion (9), we obtain the pass-through matrix for a TD model with survival function Φ_s as:

$$\Psi^{\Phi} \equiv \frac{1}{\left(\sum_{s\geq 0} \Phi_s\right)\left(\sum_{s\geq 0} \beta^s \Phi_s\right)} \begin{pmatrix} \Phi_0 & 0 & 0 & \cdots \\ \Phi_1 & \Phi_0 & 0 & \cdots \\ \Phi_2 & \Phi_1 & \Phi_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \Phi_0 & \beta\Phi_1 & \beta^2\Phi_2 & \cdots \\ 0 & \Phi_0 & \beta\Phi_1 & \cdots \\ 0 & 0 & \Phi_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{10}$$

The matrix on the right in (10) 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}. 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



(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 $\{0.2, 0.3\}$.

hazards. 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" shape is more spiked, reflecting the fact that firms adjust less in advance and more in the period of the shock itself.

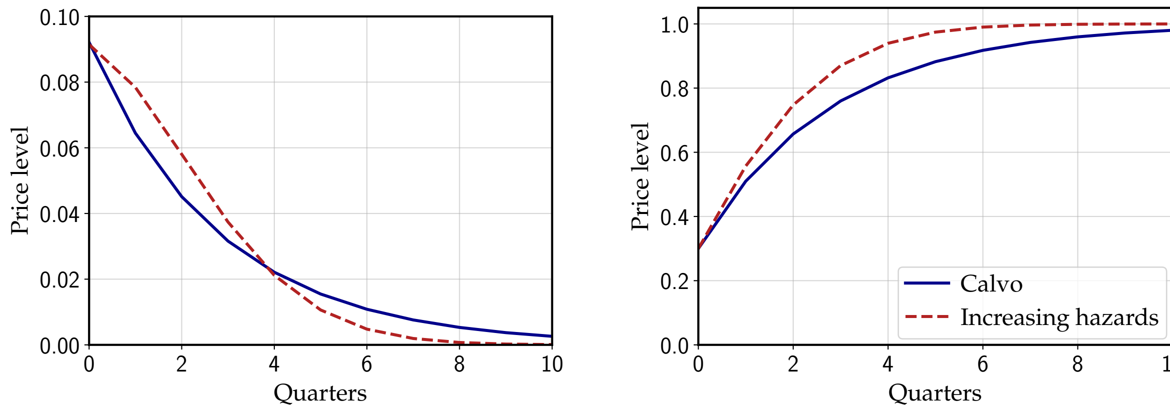
Using the expression for Ψ^{Φ} in (10), we can evaluate the impulse responses of a TD model to arbitrary marginal cost shocks. Two important special cases are that of a one-time, perfectly transitory, shock to marginal cost at date $t = 0$; and that of a permanent shock to marginal cost. For the one-time 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,

$$\text{one-time shock: } \hat{P}_t = \frac{\Phi_t}{\left(\sum_{s\geq 0} \Phi_s\right)\left(\sum_{s\geq 0} \beta^s \Phi_s\right)} \qquad \text{permanent shock: } \hat{P}_t = \frac{\sum_{s=0}^{t} \Phi_s}{\sum_{s\geq 0} \Phi_s} \tag{11}$$

We plot both impulse responses, for the Calvo model and a model with increasing hazards, in figure 2. The formulas in (11) allow the survival function Φ_t of any TD model to be read off from either impulse response. In fact, the formula for the impulse response to the permanent shock corresponds exactly to the cumulative weights that enter the law of motion (9). 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 shocks 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),



(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.

the pass-through matrix is sufficient to analyze the price level response to a shock to the money growth 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, as we demonstrate in 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 $\widehat{\mathbf{mc}} \equiv \widehat{\mathbf{MC}} - \hat{\mathbf{P}}$. Substituting this equation into (7), we can derive the price level response to a real marginal cost shock,[13]

$$\hat{\mathbf{P}} = \Psi \left(\widehat{\mathbf{mc}} + \hat{\mathbf{P}} \right) = \sum_{k \geq 1} \Psi^k \cdot \widehat{\mathbf{mc}} = \Psi \left(\mathbf{I} - \Psi \right)^{-1} \widehat{\mathbf{mc}}. \tag{12}$$

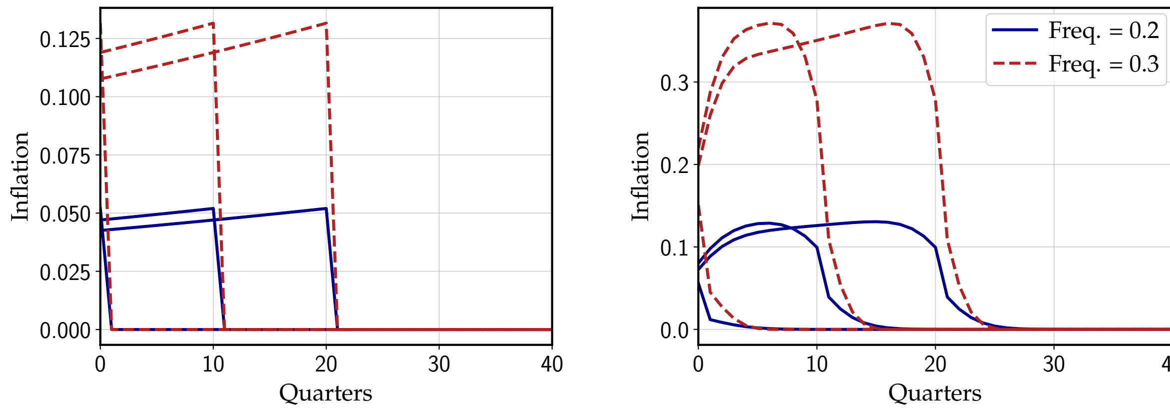
Taking first differences of (12) corresponds to left-multiplying both sides with $\mathbf{I} - \mathbf{L}$, where $\mathbf{L}$ is the lag matrix with entries of 1 one below the diagonal. We thus find the stacked inflation response

$$\pi = \left(\mathbf{I} - \mathbf{L} \right) \hat{\mathbf{P}} = \underbrace{\left(\mathbf{I} - \mathbf{L} \right) \Psi \left(\mathbf{I} - \Psi \right)^{-1}}_{\equiv \mathbf{K}} \widehat{\mathbf{mc}} \tag{13}$$

Equation (13) defines the *generalized Phillips curve* $\mathbf{K}$.[14] This matrix is the linear map from an arbitrary shock to *real* marginal cost $\widehat{\mathbf{mc}}$ to inflation. In that sense, $\mathbf{K}$ generalizes the (NK-PC) to pricing models with a general pass-through matrix Ψ. In fact, (13) describes a one-to-one mapping

[13] Although it is not immediate from (12) that the sum $\sum_{k \geq 1} \Psi^k$ converges to some bounded, finite $\Psi(I - \Psi)^{-1}$, we prove in appendix C.3 that it does so (and indeed characterize its asymptotic shape) 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 shock. The case $\theta = 1$ is relevant for the generalized Phillips curve, since shock to real cost acts like a strategic complementarity from the point of view of nominal price-setting.

[14] We pick letter $\mathbf{K}$ in order to mirror the slope parameter κ in the (NK-PC).



(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.

between the pass-through matrix Ψ and the generalized Phillips curve $\mathbf{K}$.[15]

Generalized Phillips curve for a TD model. For a general TD model, $\mathbf{K}$ can be evaluated numerically using the formula in (13), combined with the TD pass-through matrix (10). 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 (NK-PC), 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 interpretation is analogous to before. 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 some "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 don't adjust when the shock hits, as discussed by Sheedy (2010).

2.5 Calibration

To illustrate our theoretical results, and provide a benchmark for our numerical results, we will simulate two SD models whose parameterizations are inspired by Golosov and Lucas (2007)

[15]To obtain Ψ from $\mathbf{K}$, we write $\Psi = \mathbf{TK}(\mathbf{I}+\mathbf{TK})^{-1}$ where $\mathbf{T} = (T_{ts})$ with $T_{ts} = 1$ for $t \geq s$ and 0 elsewhere.

	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.

(henceforth GL) and Nakamura and Steinsson (2010) (henceforth NS). For these two models, we pick the following standard calibration.

We assume that the shock distribution $\mathcal{F}$ is normal with variance σ_ϵ^2. 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 price size of 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%. We set the discount factor to $\beta = 0.99$. Calibrated parameters are summarized in table 1.[16]

3 Exact Equivalence between SD and TD Pricing Models

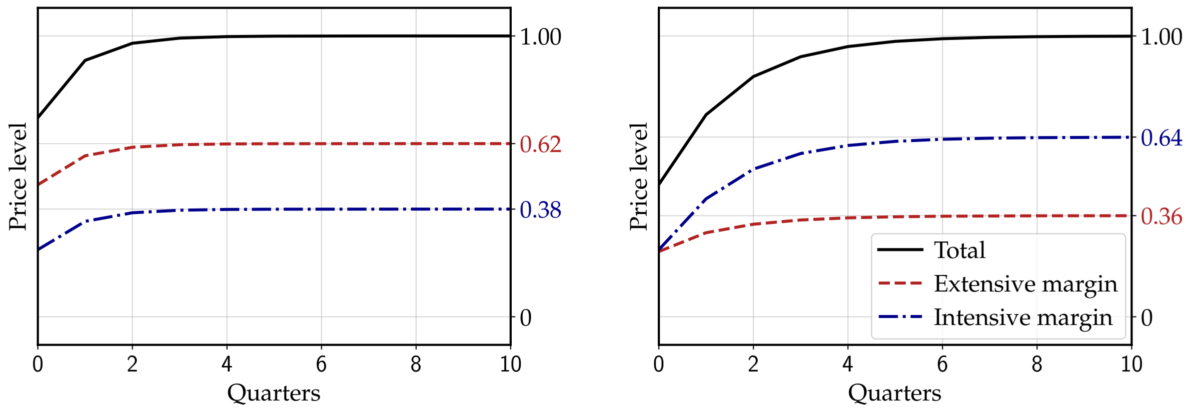
We are now ready to compare the aggregate implications of state- and time-dependent models for the dynamics of prices and inflation. Since the pass-through matrix encapsulates all the first-order implications of the pricing models introduced so far, a simple place to start is as follows. Consider an 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 given its stronger "selection effect", 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 (11) 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'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.[17] 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.

[16] Note that the homotheticity properties of our model imply that there are only two independent degrees of freedom. Given λ and ξ/σ_ϵ^2, σ_ϵ only scales the sS bands and shock size, without any implications for the aggregate behavior of prices to a first order approximation (see also Alvarez et al. 2016).

[17] As we explain in detail in section 4.4, picking Φ to match the survival function of prices implied by the SD model also does not generate the correct SD pass-through matrix.



(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.

3.1 Exact equivalence result

To make progress, we now consider the underlying drivers of the SD impulse responses in figure 4. As in the analysis in Caballero and Engel (2007), the permanent shock shifts up both the adjustment barriers $\underline{x}_t, \overline{x}_t$, and the reset price gap x_t^*. We can thus separate each total impulse response into an *extensive margin* component, driven by the shift in the adjustment barriers; and into an *intensive margin* component, driven by the shift in the reset price gap. Since the shock is evaluated to first order, this decomposition is additive.

The red and blue 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, the long-run share of the extensive margin is only 36%.

Consider repeating the TD matching strategy discussed above, 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 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,*

$$\Psi = \alpha \Psi^{\Phi^e} + (1-\alpha)\,\Psi^{\Phi^i} \tag{14}$$

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;*
- $(1-\alpha) \sum_{s=0}^{t} \Phi_s^i / \sum_{s\geq 0} \Phi_s^i$ *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 and proves that it works. 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 $1-\alpha$). 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 $(1-\alpha)\,\Psi^{\Phi^i} \cdot \widehat{\mathbf{MC}}$, and that both contributions come from TD models. For example, given equation (11), the impulse response of the SD model to a one-time shock is simply given by

$$\alpha \frac{\Phi_0^e}{\left(\sum_{s\geq 0} \Phi_s^e\right)\left(\sum_{s\geq 0} \beta^s \Phi_s^e\right)} \Phi_t^e + (1-\alpha) \frac{\Phi_0^i}{\left(\sum_{s\geq 0} \Phi_s^i\right)\left(\sum_{s\geq 0} \beta^s \Phi_s^i\right)} \Phi_t^i \tag{15}$$

Our result thus naturally generalizes the Caballero and Engel (2007) decomposition to arbitrary shocks and to the entire impulse response.[18] Since the pass-through matrix characterizes the entire 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 follows 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 (14) to be of size $T \times T$. From (10), 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, (14) effectively is a reduction from T^2 dimensions down to $2T-1$ dimensions.[19]

[18] 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.

[19] $2T-2$ 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 (14).

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—in fact, this is why this exercise is the norm among 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, one computes the impulse response to a permanent shock. This response is then split up into components coming from only shifting the sS bands vs. only shifting the reset price gap. From those components, one gets Φ^e, Φ^i, α as outlined above, and computes the right hand side of (14) using the formula in (10). This gives the pass-through matrix Ψ for an SD model. In turn, this allows one to simulate arbitrary shocks to nominal costs by taking the matrix product of Ψ and the shock vector. In addition, using (13), we can construct the generalized Phillips curve $\mathbf{K}$ and simulate arbitrary shocks to real marginal cost shocks, as well.

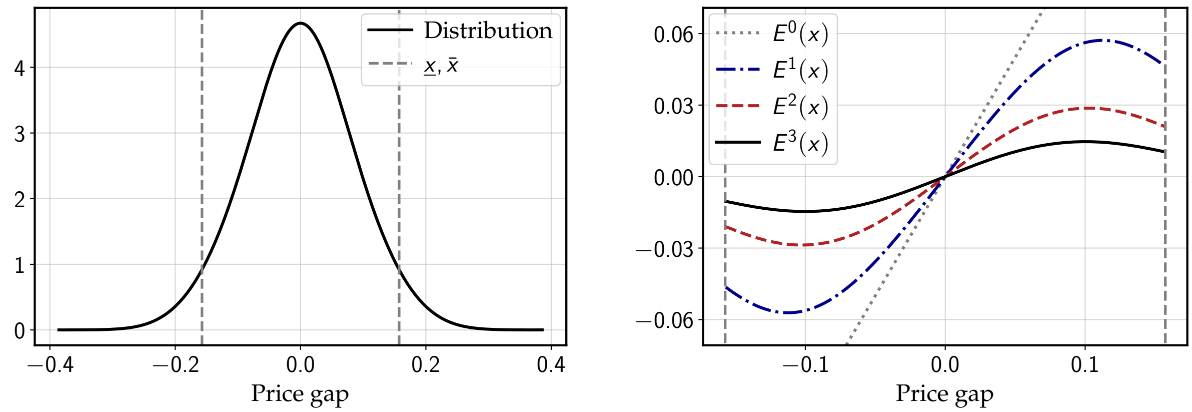
Continuous time version of our result. While our result is set in discrete time, a similar result holds in continuous time. We present it in appendix A.

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 these policy changes are determined in response to future marginal cost shocks (the policy equation). In both cases, we find a central role for $E^t(x)$. When we combine the law of motion and policy equation to obtain the pass-through matrix, we see that this matrix is a weighted sum of two terms—representing the extensive and intensive margins—each of which has exactly the same form as in the time-dependent case (10), with survival functions derived from $E^t(x)$. We conclude by identifying these terms with the decomposition (14) in proposition 1.

Expected price gaps. Consider a menu cost model in steady state, with $\log MC_t = 0$. 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 $E^t(x) \equiv \mathbb{E}_0\left[x_{it}|x_{i0} = x\right]$ as the firm's expected price gap at the end of period t.[20] 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 $E^t(0) = 0$.

[20] 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.

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

$$E^t(x) = (1-\lambda) \int_{-\bar{x}}^{\bar{x}} f(x'-x) E^{t-1}(x') dx', \tag{16}$$

taking expectations over $E^{t-1}(x')$ using the no-adjustment transition probability from x to x'. (The contribution from price adjustments to (16) 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, 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 left panel in figure 5 plots the stationary distribution $g(x)$ of price gaps after shocks have realized, but before firms have adjusted their price. Prices that lie outside the steady state sS band $[\underline{x}, \overline{x}]$, and a random share λ of prices inside the bands, adjust to 0.

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

$$\log P_t = (1-\lambda) \int_{\underline{x}_t}^{\overline{x}_t} x g_t(x) dx + \left(\lambda + (1-\lambda) \left(1 - \int_{\underline{x}_t}^{\overline{x}_t} g_t(x) dx \right) \right) x_t^* \tag{17}$$

where $g_t(x)$ is the pre-adjustment density, and the first and second terms are the contributions from adjusters and non-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. We can then rewrite $\log P_t$ as

$$\log P_t = (1-\lambda)\int_{\underline{x}_{t-s}}^{\overline{x}_{t-s}} E^s(x)g(x)dx + \left(\lambda + (1-\lambda)\left(1 - \int_{\underline{x}_{t-s}}^{\overline{x}_{t-s}} g(x)dx\right)\right) E^s(x_t^*) \tag{18}$$

where we obtain the average value of price gaps x at date t by taking the average over the expected price gap $E^s(x)$ at date $t-s$.

Totally differentiating (18) around the steady state, we have

$$d\log P_t = (1-\lambda)E^s(\overline{x})g(\overline{x})(d\overline{x}_{t-s} + d\underline{x}_{t-s}) + \text{freq}\cdot(E^s)'(0)dx_{t-s}^* \tag{19}$$

where the first term simplifies due to symmetry, the second term simplifies since $E^s(0) = 0$, and $\text{freq} = \lambda + (1-\lambda)\left(1 - \int_{-\overline{x}}^{\overline{x}} g(x)dx\right)$ is the steady-state price adjustment frequency.

Equation (19) gives the first-order response, around the steady state, of $\log P_t$ to changes in policy 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

$$d\log P_t = (1-\lambda)g(\overline{x})\sum_{s=0}^{t} E^s(\overline{x})(d\overline{x}_{t-s} + d\underline{x}_{t-s}) + \text{freq}\cdot\sum_{s=0}^{t}(E^s)'(0)dx_{t-s}^*. \tag{20}$$

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.

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 (2) around the steady state, this implies $dV_s(x) = -x$. 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

$$dV_t(x) = \beta(1-\lambda)\int_{\underline{x}}^{\overline{x}} f(x'-x)dV_{t+1}(x')dx \tag{21}$$

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.[21]

Given that $dV_s(x) = E^0(x) = -x$, and that (21) has exactly the same form as our earlier recursion (16), except with an additional discount factor β, It follows that

$$dV_t(x) = -\beta^{s-t}E^{s-t}(x). \tag{22}$$

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 $V_t(\overline{x}_t) = V_t(\underline{x}_t) = V_t(x_t^*) + \zeta$, 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 $d\overline{x}_t = -\left(dV_t(\overline{x}) - dV_t(0)\right)/V'(\overline{x})$,

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

$d\underline{x}_t = -\left(dV_t(\underline{x}) - dV_t(0)\right)/V'(\underline{x})$, and $dx_t^* = -dV_t'(0)/V''(0)$, which combined with our results above become $d\overline{x}_t = d\underline{x}_t = \frac{\beta^{s-t}E^{s-t}(\overline{x})}{\sum_{u=0}^{\infty}\beta^u E^u(\overline{x})} d\log MC_s$ 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

$$d\overline{x}_t = d\underline{x}_t = \frac{\sum_{s\geq t}\beta^{s-t}E^{s-t}(\overline{x})d\log MC_s}{\sum_{s\geq t}\beta^{s-t}E^{s-t}(\overline{x})} \tag{23}$$

$$dx_t^* = \frac{\sum_{s\geq t}\beta^{s-t}(E^{s-t})'(0)d\log MC_s}{\sum_{s\geq t}\beta^{s-t}(E^{s-t})'(0)} \tag{24}$$

i.e. that both the changes in thresholds $d\overline{x}_t, d\underline{x}_t$ 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 (20), assuming $d\overline{\mathbf{x}} = d\underline{\mathbf{x}}$, as

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

where $\overline{\mathbf{x}} \equiv (\overline{x}_0, \overline{x}_1, \overline{x}_2, \ldots)'$, etc. Then, substituting in the vector form of (23)–(24) and rearranging, this becomes

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

$$+\frac{\text{freq}\cdot\sum_{s\geq 0}E^{s\prime}(0)}{\left(\sum_{s\geq 0}E^{s\prime}\right)\left(\sum_{s\geq 0}\beta^s E^{s\prime}(0)\right)}\begin{pmatrix} E^{0\prime}(0) & 0 & 0 & \cdots \\ E^{1\prime}(0) & E^{0\prime}(0) & 0 & \cdots \\ E^{2\prime}(0) & E^{1\prime}(0) & E^{0\prime}(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} E^{0\prime}(0) & \beta E^{1\prime}(0) & \beta^2 E^{2\prime}(0) & \cdots \\ 0 & E^{0\prime}(0) & \beta E^{1\prime}(0) & \cdots \\ 0 & 0 & E^{0\prime}(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\widehat{\mathbf{MC}}$$

We see that each term above, aside from the numerators, has *exactly* the form of the time-dependent pass-through matrix (10). 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.[22] Hence, if we define the

[22] Also, the first entries in these sequences, $\frac{E^0(\overline{x})}{\overline{x}}$ and $E^{0\prime}(0)$, equal 1, as required for a survival function.

survival functions $\Phi^e_s \equiv \frac{E^s(\bar{x})}{\bar{x}}$ and $\Phi^i_s \equiv E^{s\prime}(0)$, this pass-through matrix simplifies to just

$$\Psi = 2(1-\lambda)g(\bar{x})\left(\sum_{s\geq 0}\frac{E^s(\bar{x})}{\bar{x}}\right)\Psi^{\Phi^e} + \text{freq}\left(\sum_{s\geq 0}E^{s\prime}(0)\right)\Psi^{\Phi^i}, \tag{25}$$

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 $\bar{x}_t$ 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 (10) 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.[23] For (25) be consistent with this, the sum of the coefficients on Ψ^{Φ^e} and Ψ^{Φ^i} must equal 1. Hence defining $\alpha \equiv 2(1-\lambda)g(\bar{x})\left(\sum_{s\geq 0}\frac{E^s(\bar{x})}{\bar{x}}\right)$, we can rewrite (25) as just

$$\Psi = \alpha\Psi^{\Phi^e} + (1-\alpha)\Psi^{\Phi^i} \tag{26}$$

which is identical to (14) in proposition 1. Consistent with proposition 1, the terms $\alpha\Psi^{\Phi^e}$ and $(1-\alpha)\Psi^{\Phi^i}$ give the responses when only the extensive or intensive margins, respectively, adjust.

The expressions here 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^e_t, Φ^i_t and α from these formulas, and therefore form the pass-through matrix from this information alone—without ever needing to solve the full model. We follow a related approach in section 6.

3.3 Properties of the equivalent time-dependent models

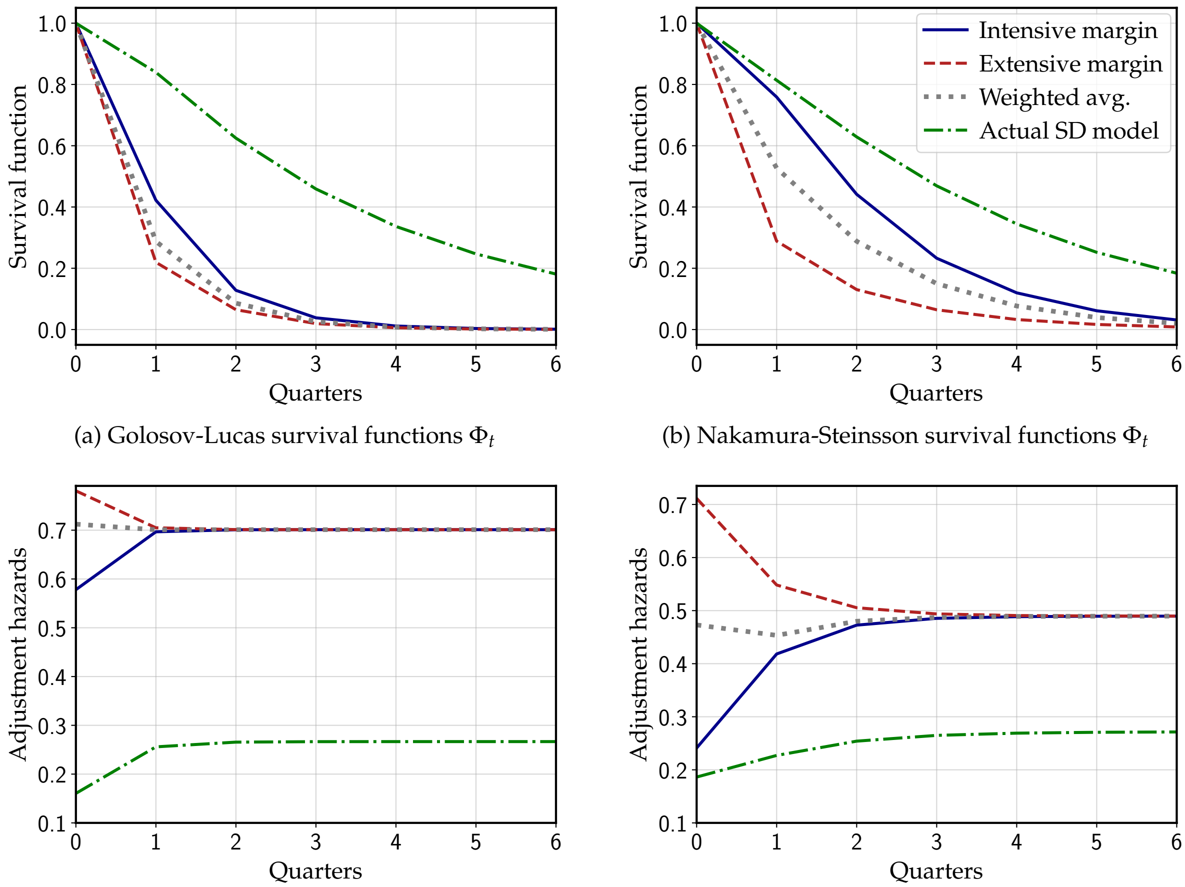
At the heart of the equivalence result in proposition 1 are the virtual survival functions Φ^e_t, Φ^i_t. Here, we study these functions in our calibrated examples and discuss their general properties.

Figure 6 plots Φ^e_t, Φ^i_t 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.

It is tempting to compare these 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 $s \leq t$,

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

[23] It follows immediately from (23)–(24) 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.



(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 $\alpha\Phi_t^e + (1-\lambda)\,\Phi_t^i$, with corresponding adjustment hazards λ_t for the GL and NS models, calibrated as in table 1.

The actual SD hazards in both the GL and NS are increasing (see e.g. Alvarez, Borovičková and Shimer 2021) 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 hazard of the aggregate price level in response to a permanent nominal marginal cost shock is strictly below 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.[24]

Following Alvarez and Lippi (2022), we can attribute the gap between the asymptotic adjustment hazard and prices and actual hazard to a selection effect. Our analytical expressions for the extensive and intensive margins shed light on selection effects on both the extensive and intensive margin. Recall from equation (**??**) that the extensive margin virtual survival curve Φ_t^e is given by $E^t(\overline{x})/\overline{x}$, 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:

$$\Phi_t^e = \Phi_t^{\text{actual}} \times \frac{\mathbb{P}\left(\text{no adj. until } t | x_0 = \overline{x}\right)}{\mathbb{P}\left(\text{no adj. until } t | x_0 = 0\right)} \times \frac{\mathbb{E}\left[x_t | \text{no adj. until } t, x_0 = \overline{x}\right]}{\overline{x}} \tag{27}$$

The two factors on the right of (27) give the two reasons why Φ_t^e declines faster than Φ_t^{actual}. First, prices at the boundary $\overline{x}$ 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 (27) 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 (27) shows that there is also a selection effect in the intensive margin of price adjustment. Indeed, equation (**??**) shows that we have:

$$\Phi_t^i = \Phi_t^{\text{actual}} \times \left.\frac{\partial}{\partial x}\mathbb{E}\left[x_t | \text{no adj. until } t, x_0 = x\right]\right|_{x=0} \tag{28}$$

The right factor in (28) 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 the middle of the sS band. Hence the marginal effect of setting a higher price today on the future price gap is attenuated by selection, which compresses the surviving price gaps.

Indeed, as proposition 2 shows, asymptotically the extensive and intensive selection effects are equally powerful, leading to the same hazard $\lambda_\infty^{virtual} > \lambda_\infty^{actual}$. 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 (27) below 1, but also makes the selection effect in the rightmost factor of (27) initially much stronger than in (28). 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

[24] 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.

shocks we study in figure 9, as well as for the leptokurtic shocks discussed in appendix D.2. 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.[25]

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.

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 $M \geq 2\overline{x}$.[26]

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,

$$E^1(x) = \underbrace{(1-\eta)\,x}_{\text{zero shock}} + \underbrace{\eta \frac{2\overline{x}}{M} \cdot 0}_{\text{uniform shock, no adj.}} + \underbrace{\eta\left(1 - \frac{2\overline{x}}{M}\right) \cdot 0}_{\text{uniform shock, adj.}} = (1-\eta)\,x$$

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 in (**??**) and (**??**), we therefore obtain:

$$\Phi^i_t = \Phi^e_t = (1-\eta)^t \qquad \text{and} \qquad \alpha = \frac{2\overline{x}}{M} \tag{29}$$

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.

Gertler and Leahy (2008) pointed out that this virtual Calvo frequency is higher than the actual frequency of price adjustment, $\text{freq} = \eta\left(1 - \frac{2\overline{x}}{M}\right)$,[27] delivering less monetary non-neutrality in the menu cost model than would be inferred from the frequency of price adjustment alone. Equation

[25] A counterexample can be obtained by setting $\bar{x} = 1$ and an idiosyncratic shocks p.d.f. $f(x) \propto e^{-(x/1.5)^8}$. The extensive and intensive margins of this model are then both non-monotonic.

[26] Note that this distribution of idiosyncratic shocks does not satisfy the regularity conditions in section 2.1. We show in the appendix, however, that the exact equivalence result still holds for this case.

[27] Note that this expression for the frequency verifies the expression for $1-\alpha$ in footnote **??**.

(29) shows that the gap between the actual and the virtual frequency $\frac{2\bar{x}}{M}$, is equal to the weight on the extensive margin α. This is intuitive, since this weight is a measure of the importance of selection.[28]

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.4.

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 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 grey line, is, in fact, approximately constant in these examples. This suggests that these models may effectively be close to a Calvo model in practice.

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.

4.1 Distance between pricing models

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

$$\text{dist}\left(\mathbf{J}, \mathbf{J}'\right) = \frac{\|\mathbf{J} - \mathbf{J}'\|}{\|\mathbf{J}\|} \tag{30}$$

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 (30) captures the average slope of the Phillips curve: how much of an inflation response a unit-standard-deviation real marginal cost shock can

[28]In this example, the weight on the extensive margin α is exactly equal to the gap between the virtual and the actual asymptotic hazard. In simulations of more general models we have found the two metrics to still be correlated. In these cases, the latter provides a more useful measure of selection that the former, since it directly relates to the difference between actual and measured adjustment probability.

generate. For a Calvo model, we have[29]

$$\|\mathbf{K}^{\text{Calvo}}(\kappa)\| = \sup_{\widehat{\mathbf{mc}}:\, \|\widehat{\mathbf{mc}}\|=1} \|\mathbf{K}^{\text{Calvo}}(\kappa) \cdot \widehat{\mathbf{mc}}\| = \frac{\kappa}{1-\beta} \tag{31}$$

The numerator in (30) captures the worst-case standard deviation of the differential inflation response across the two models,

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

One can evaluate this norm similarly to (31), 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

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

Intuitively, our measure of distance in (30) captures the relative difference in Phillips curve slopes. In the following, we apply (30) 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, $\mathbf{J}' = \mathbf{K}^{\text{Calvo}}(\kappa)$.

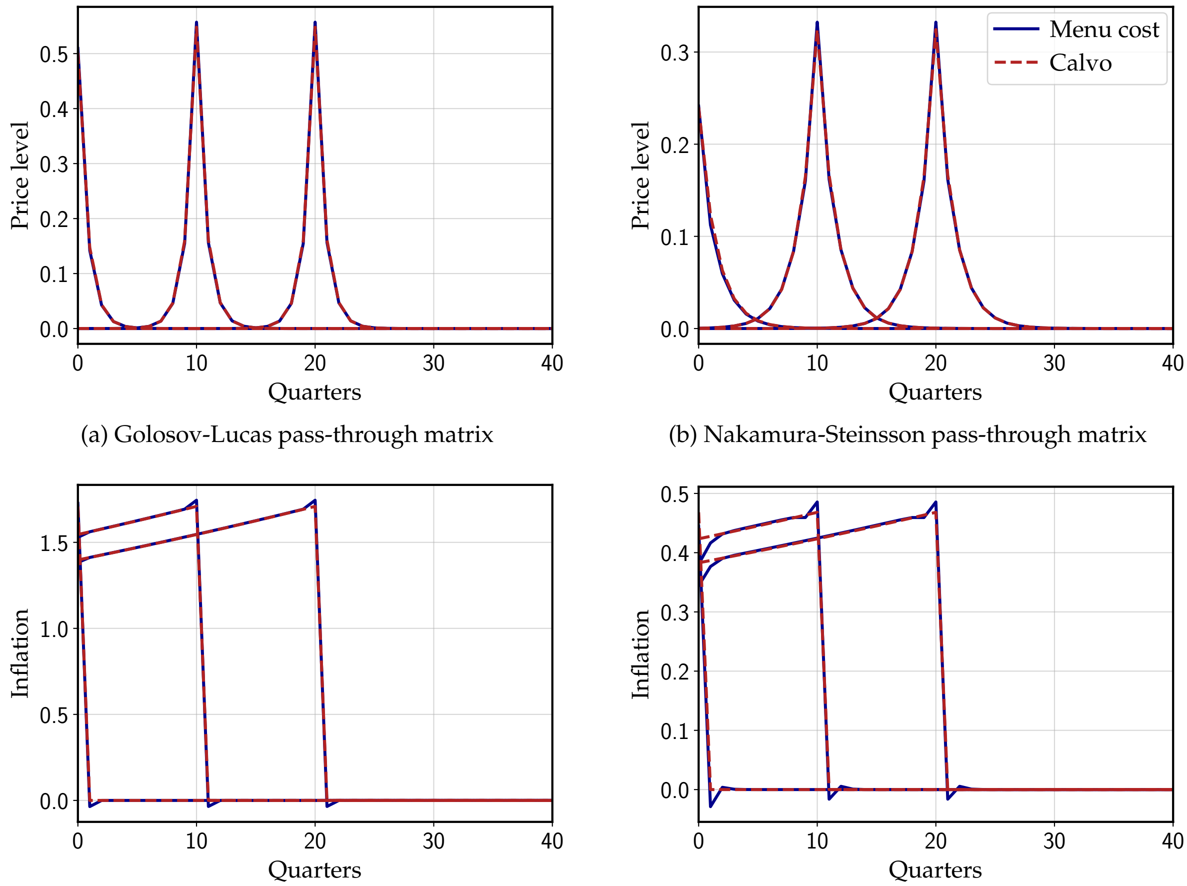
In principle, our distance measure (30) can also be used to compare pass-through matrices. Since pass through matrices and generalized Phillips curves are related by the one-to-one mapping in (13), the distance measures end up being very similar. However, the distance between Phillips curve matrices 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 (30) 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

[29]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 (NK-PC), we have $\|\boldsymbol{\pi}\| \leq \kappa\|\widehat{\mathbf{mc}}\| + \beta\|\boldsymbol{\pi}\|$ and thus $\|\boldsymbol{\pi}\|/\|\widehat{\mathbf{mc}}\| \leq \frac{\kappa}{1-\beta}$.



(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 (30) applied to Phillips curve matrices.

0.707 and 0.487. Given $\beta = 0.99$, this implies Calvo slopes κ of 1.709 and 0.468.

Remark 2. The GL and NS models are numerically 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 et al. 2021, 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 due to the difference between virtual and actual hazards. As figure 6 shows, in both the GL and the NS models, actual hazards increase, but effective virtual hazards do not. In other words, the selection effect undoes the inertia in the Phillips curve that

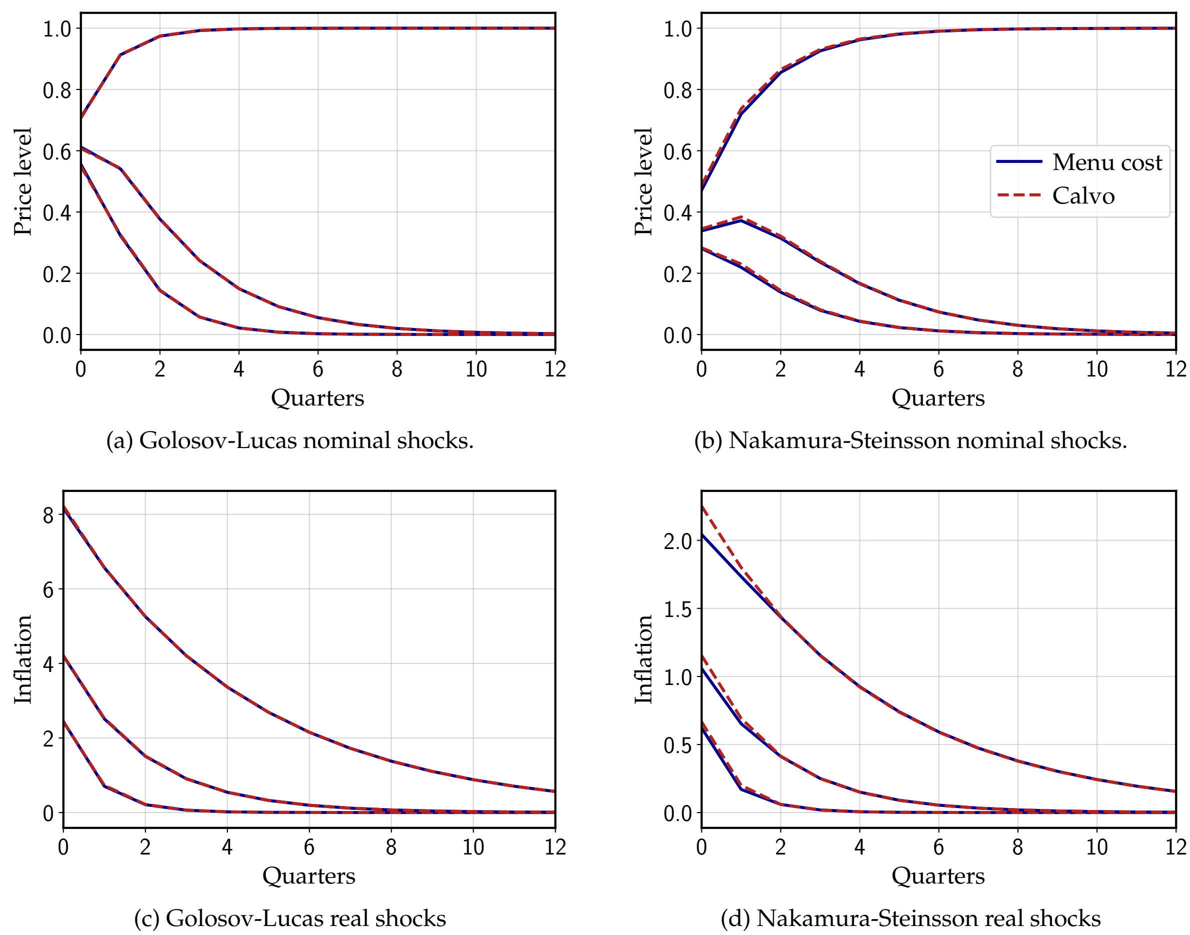


Figure 8: Price level responses to AR(1) marginal cost shocks.

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 (30). Shock persistence values are $\{0.3, 0.6, 1\}$ for nominal shocks (top) and $\{0.3, 0.6, 0.8\}$ for real shocks (bottom).

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.

By contrast, our numerical equivalence result establishes that the entire impulse response to *arbitrary 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 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 higher prevalence of free adjustments. The reason is as follows. While both models feature the same frequency of price changes, in the GL model, these adjustments are entirely triggered by price gaps leaving the inaction region. This leads to faster mixing of price gaps, and hence faster convergence of the intensive and extensive margin hazards relative to the NS model, as is clear from figure 6. In turn, this faster convergence makes GL more and NS less "Calvo-like".

Estimating the NK-PC 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 (NK-PC), 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 can 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.

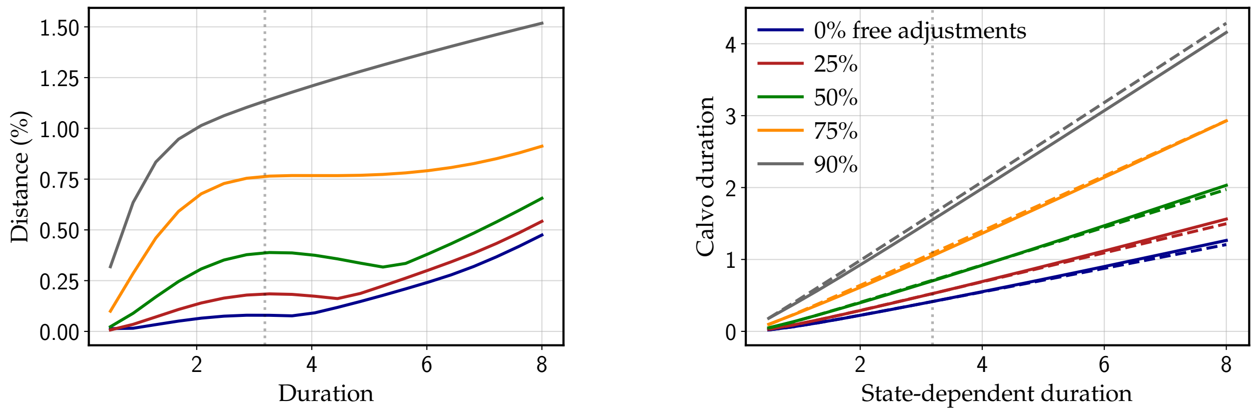
4.3 Robustness to parameterization of the canonical menu cost model

The GL and NS models are only two parametrizations of the canonical menu cost model. Here, we explore systematically the two-dimensional parameter space of this model within the class of normal idiosyncratic shocks.[30]

Figure 9(a) plots the relative distance (30) 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 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

[30] In appendix D.2 we consider an extension with leptokurtic shocks.



(a) Distance for various duration and share of free adjustments values.

(b) Distance-minimizing (solid) and ALL-implied (dashed) Calvo duration.

Figure 9: General parameter values.

Note: the left panel shows the distance between the state-dependent Phillips curve and the best-fitting Calvo approximation, measured with distance (30). The right panel shows the duration of the Calvo minimizer, as well as the one implied by the ALL 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 share of free adjustment under consideration. . The vertical dotted lines show the empirical duration that we target in our baseline calibration.

increase the distance to Calvo. The distance only falls to zero for adjustment shares very close to 100% (not shown).

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

In this section, we explain the Calvo model appears to approximate the menu cost model so well.

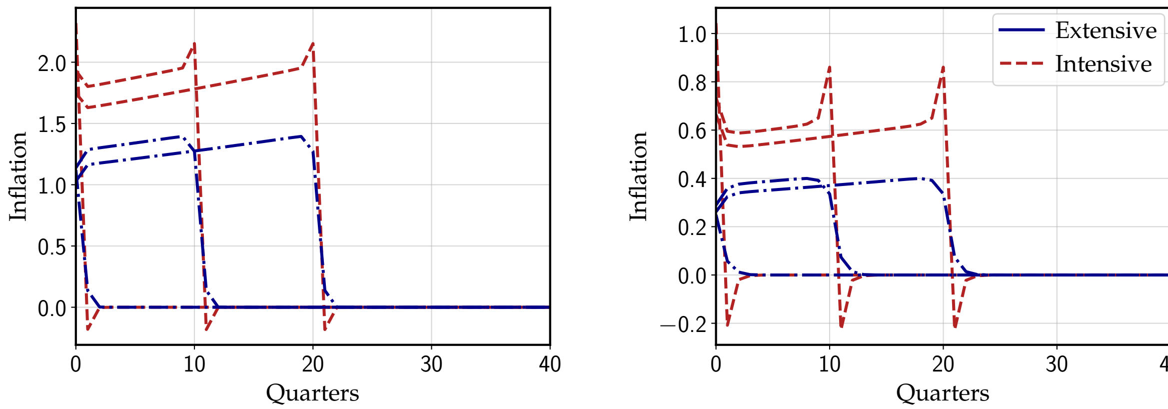
In our GL and NS parameterizations of the canonical menu cost model, we traced back the close fit between the Calvo model and the SD models to the fact that underlying extensive and intensive margin hazards evolved in opposite directions and converged to the same limit, per proposition 2 and remark 1. In particular, we noted that the average survival function, $\alpha\Phi_t^e + (1-\alpha)\,\Phi_t^i$, was close to a Calvo survival function $\Phi_t^{calvo} = (1-\lambda)^t$.

One reason why the average survival function is relevant is that, if two models have the same average survival function, then they give the same impulse response to a permanent shock. Indeed, proposition 1 together with (11) implies that the response to a permanent shock of a canonical SD model is always

$$\widehat{P}_t = \frac{\sum_{s=0}^{t}\left(\alpha\Phi_s^e + (1-\alpha)\,\Phi_s^i\right)}{\sum_{s\geq 0}\left(\alpha\Phi_s^e + (1-\alpha)\,\Phi_s^i\right)}$$

Hence, all that is relevant for this impulse response is the average survival $\alpha\Phi_s^e + (1-\alpha)\,\Phi_s^i$.

Beyond the permanent shock case, the average survival function is not sufficient to pin down the impulse responses to other shocks. Suppose, for instance, that two TD models average exactly to a Calvo, $\alpha\Phi_t^e + (1-\alpha)\,\Phi_t^i = (1-\lambda)^t$, but are not Calvo themselves. Then, equation (15) shows that they will not deliver the same response to a one-time shock; and the logic holds more gener-



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

Figure 10: Decomposing the generalized Phillips curve into intensive and extensive margin

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

ally. However, if these two models are themselves "close" to Calvo so that $\Phi_t^e = (1-\lambda)^t + \eta_t$ with η_t small, then to first order in η_t, all nominal impulse responses are the same as Calvo impulse responses. Through the mapping in (13), this then implies that the generalized Phillips curves are themselves close. Since, in figure 6, the average virtual survival of the GL model is extremely close to flat, this explains why the Calvo model does so well for the generalized Phillips curve.

Figure 6 shows that the average virtual survival of the NS model, on the other hand, differs from that of a Calvo. To understand the implications of this for the generalized Phillips curve, we consider a situation where the survivals are most different from Calvo, because there is either only an intensive margin or only an extensive margin. Figure 10 reports the generalized Phillips curve for the GL and NS model in this case. We see that these two models separately produce fairly different GPC. In particular, there is persistence in the extensive margin and anti-persistence in the intensive margin. This difference is most pronounced for the NS model.

A striking feature, however, is that the shape is still fairly similar: 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 any price-setting model with any pass-through matrix.

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 has several implications. First, it shows that the shape of the generalized Phillips curve will always look similar to Calvo in that it will always feature β discounting of anticipated shocks in any model. Second, it shows that the extreme "forward-lookingness" of the (NK-PC), with future shocks discounted at rate β irrespective of the horizon—which plays an major role in the forward guidance puzzle (Del Negro, Giannoni and Patterson, 2013)—is present in all price-setting models we consider in this paper.

4.5 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 completely side-step the need for 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,

$$\kappa = \frac{\text{freq}\,(1-\beta\,(1-\text{freq}))}{1-\text{freq}} = \frac{1-\beta\frac{d}{1+d}}{d} \tag{32}$$

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,

$$\frac{\text{Kur}^{\text{Calvo}}}{\text{freq}^{\text{Calvo}}} \simeq \frac{\text{Kur}^{\text{MC}}}{\text{freq}^{\text{MC}}} \tag{33}$$

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, (33) 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.[31]

In conclusion, equations (32) and (33), combined with the relationship between Calvo kurtosis and frequency, allow us to obtain a Phillips curve based entirely on the ratio of kurtosis to

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

frequency in the menu cost model.[32] In the limit as $\beta \to 1$, this becomes:

$$\kappa \simeq \frac{4}{\left(\frac{1}{3}\frac{\text{Kur}^{MC}}{\text{freq}^{MC}}\right)^2 - 1} \tag{34}$$

The quantitative implications of (34) 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 to 4. Implementing (34) with these values 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 (34), 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.

4.6 Additional robustness exercises

In appendix D.2, we show that the numerical equivalence result between the canonical model and Calvo appears to be quite robust by considering several extensions to our numerical analysis: (a) we allow for leptokurtic shocks; (b) we allow for two products as in Midrigan (2011);[33] (c) we allow for multiple sectors as in Nakamura and Steinsson (2010); and (d) we compare SD models and Calvo models for large (non-infinitesimal) shocks.

In a final extension, we consider steady-state inflation, which in menu cost models affects the frequency of price change. When we recalibrate the model to match the same data frequency at different rates of inflation, we obtain a good fit to almost exactly the same Calvo parameter. When we do not recalibrate, and instead allow the steady-state frequency of price change to vary with inflation, we continue to find a good fit to Calvo, but the Phillips curve slope κ increases with steady-state inflation (see figure D.3).

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 (eg, Alvarez and Lippi 2014).

We have also studied the generalized Phillips curve as solution to a particular fixed point, solving for inflation as a function of real marginal cost in (13), by analogy to the New Keynesian

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

[33] More products would make the model more similar to a Taylor model, as shown by Alvarez et al. (2016), and would therefore make the numerical fit of a Calvo model worse.

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.[34] 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 show how the result extends to a more complex DSGE model.[35]

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 the household are the usual ones,

$$C_t^{-\sigma} = \frac{P_t}{P_{t+1}} (1 + i_t) C_{t+1}^{-\sigma} \qquad \text{and} \qquad b N_t^{\varphi} = \frac{W_t}{P_t} C_t^{-\sigma} \tag{35}$$

Consumption C_t is an index bundling many varieties i,

$$C_t \equiv \left(\int_0^1 \left(\frac{C_{it}}{A_{it}} \right)^{\frac{\zeta-1}{\zeta}} di \right)^{\frac{\zeta}{\zeta-1}} \tag{36}$$

where $\zeta > 1$ is the elasticity of substitution between varieties; and A_{it} are idiosyncratic preference shifters. Aggregate output is simply equal to consumption, $Y_t = C_t$, in the model. Demand for

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

[35] As pointed out by Fernández-Villaverde (2010, p.40), GE 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."

variety i and the aggregate price index are given by

$$Y_{it} = A_{it}^{1-\zeta} \left(\frac{P_{it}}{P_t} \right)^{-\zeta} Y_t \qquad \text{and} \qquad P_t = \left(\int_0^1 (A_{it} P_{it})^{1-\zeta} \, di \right)^{\frac{1}{1-\zeta}} \tag{37}$$

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} is an iid standard normal random variable. Firm i's real profits at date t are given by

$$\Pi_{it} = \frac{P_{it}}{P_t} Y_{it} - \frac{W_t}{P_t} N_{it} = \left(\frac{P_{it}}{P_t} - \frac{W_t}{P_t} \frac{1}{A_{it}} \right) \cdot A_{it}^{1-\zeta} \left(\frac{P_{it}}{P_t} \right)^{-\zeta} Y_t \tag{38}$$

The firm's statically optimal price P_{it}^* is therefore given by the usual constant markup rule

$$P_{it}^* = \frac{\zeta}{\zeta - 1} \frac{W_t}{A_{it}} \tag{39}$$

Substituting out A_{it} from (38) using (39), we can express profits just in terms of the ratio of the actual price P_{it} to the statically optimal price P_{it}^*,

$$\Pi_{it} = \left(\frac{\zeta}{\zeta - 1} \frac{W_t}{P_t} \right)^{1-\zeta} Y_t \cdot \left(\left(\frac{P_{it}}{P_{it}^*} \right)^{1-\zeta} - \frac{\zeta - 1}{\zeta} \left(\frac{P_{it}}{P_{it}^*} \right)^{-\zeta} \right) \tag{40}$$

As before, we define firm i's price gap x_{it} as the log-difference between a firm's price and its optimal price, excluding the nominal wage

$$x_{it} = \log P_{it} - \log \left(\frac{\zeta}{\zeta - 1} \frac{1}{A_{it}} \right)$$

With this notation, profits can be written entirely as a function of the price gap and aggregate variables,[36]

$$\Pi_{it} = \left(\frac{\zeta}{\zeta - 1} \frac{W_t}{P_t} \right)^{1-\zeta} Y_t \cdot \left(e^{(1-\zeta)(x_{it} - \log W_t)} - \frac{\zeta - 1}{\zeta} e^{-\zeta (x_{it} - \log W_t)} \right) \equiv Z_t \cdot F\left(x_{it} - \log W_t \right)$$

where we collect the aggregate profit shifters in $Z_t \equiv \left(\frac{\zeta}{\zeta-1} \frac{W_t}{P_t} \right)^{1-\zeta} Y_t$ and write the idiosyncratic component with function $F(x) \equiv e^{(1-\zeta)x} - \frac{\zeta-1}{\zeta} e^{-\zeta x}$. F has a local maximum at 0, that is, $F'(0) = 0$,

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

$F''(0) < 0$. We can also express the price level P_t from (37) in terms of price gaps,

$$P_t = \frac{\zeta}{\zeta - 1} \left(\int_0^1 e^{(1-\zeta)x_{it}} di \right)^{\frac{1}{1-\zeta}} \tag{41}$$

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

As in section 2, we assume firms have to pay a random menu cost $\xi_{it} \in \{0, \xi\}$ when changing their prices, where as before, the probability of a free adjustment, $\xi_{it} = 0$, is parametrized by $\lambda \in [0, 1)$, and $\xi > 0$. 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. As we show below, this allows us to consider the limit of small σ_ϵ. Given this, firm i's profit maximization problem reads

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t C_t^{-\sigma} \left[Z_t \cdot F\left(x_{it} - \log W_t\right) + \sigma_\epsilon^2 \xi_{it} \frac{W_t}{P_t} 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}} \right] \tag{42}$$

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

$$\Xi_t \equiv \int_0^1 \sigma_\epsilon^2 \xi_{it} 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}} di \tag{43}$$

Aggregate profits $\Pi_t \equiv \int_0^1 \Pi_{it} di$ are paid directly to the representative agent. With the help of (41), aggregate labor demand by firms can be written as

$$N_t^d \equiv Y_t \Delta_t + \Xi_t \tag{44}$$

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.

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 (35) 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 $W_{ss} = 1$.[37]

In the case with no idiosyncratic shocks or menu costs ($\sigma_\epsilon = 0$), then all price gaps are zero, and we also have $\Delta_{ss} = 1$, $\Xi_{ss} = 0$, $P_{ss} = \frac{\zeta}{\zeta-1}W_{ss}$, $Y_{ss} = N_{ss}$, 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, 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.

Proposition 4. *Fix an arbitrary perturbation $\{d\nu_t\}$ to the intercept of the Taylor rule, and consider the resulting $\{dY_t(\sigma_\epsilon)\}$, $\{d\pi_t(\sigma_\epsilon)\}$, and $\{di_t(\sigma_\epsilon)\}$ in the aggregate perturbation solution, as a function of the scale of idiosyncratic risk σ_ϵ. Define $\hat{Y}_t \equiv dY_t/Y_{ss}$, $\hat{\pi}_t \equiv d\pi_t$, $\hat{i}_t = di/(1+i_{ss})$, and $\hat{\nu}_t \equiv d\nu_t$, with ss subscripts referring to the steady state with no idiosyncratic or aggregate risk. As $\sigma_\epsilon \to 0$, these approach the solution of the system*

$$\hat{Y}_t = \hat{Y}_{t+1} - \frac{1}{\sigma}(\hat{i}_t - \hat{\pi}_{t+1}) \tag{45}$$

$$\hat{\boldsymbol{\pi}} = (\varphi+\sigma)\,\mathbf{K}\cdot\hat{\mathbf{Y}} \tag{46}$$

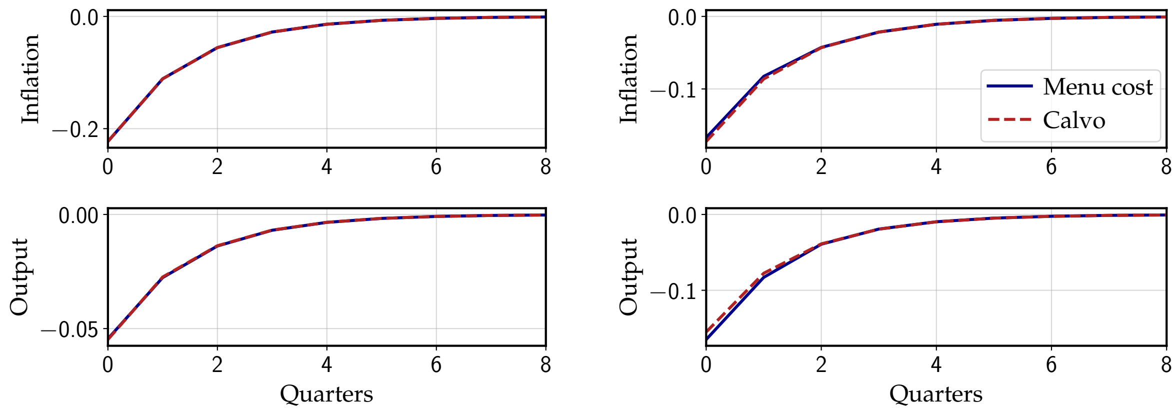
$$\hat{i}_t = \phi\hat{\pi}_t + \hat{\nu}_t \tag{47}$$

where $\mathbf{K}$ is the generalized Phillips curve implied by the canonical menu cost model in section 2.1, given the same distribution of ϵ, and replacing the menu cost ζ by $\frac{W_{ss}}{Z_{ss}P_{ss}}\frac{2\zeta}{F''(0)}$.

Note that (45)–(47) is the standard linearized three-equation New Keynesian model, except 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 (46). In short, for small enough idiosyncratic risk, the entire pricing side of the model can be summarized by (46) for the purpose of characterizing first-order aggregate impulse responses.

Figure 11 implements (45)–(47), 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

[37]The remaining equations characterizing 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.



(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 (45) — (47).

are closely matched when (46) is replaced by the New Keynesian Phillips curve for the best-fitting Calvo approximation found in section 4.

In appendix E, we show that (45)–(47) continue to provide an excellent approximation even when σ_ϵ is set to 1, so that the distribution of shocks ϵ to the price target is the same as in our original calibration, which targeted 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.

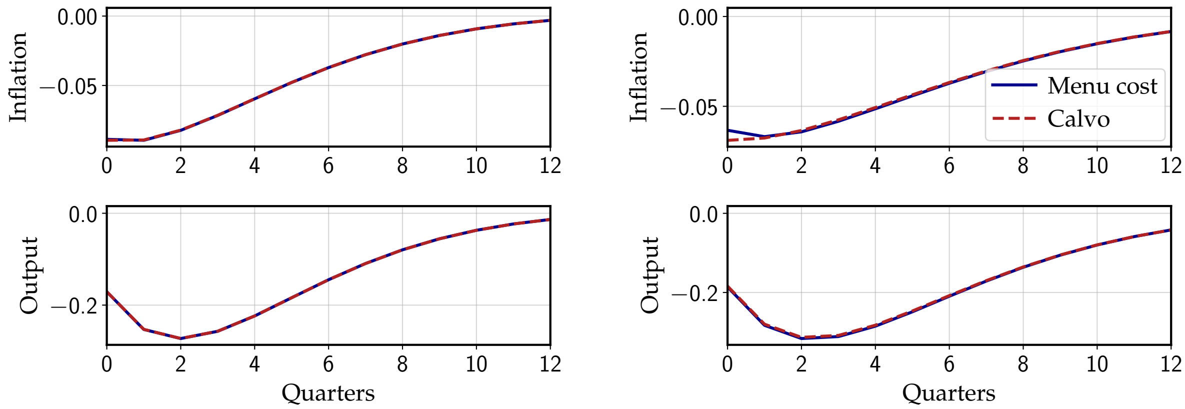
Following Nakamura and Steinsson (2010), we now introduce strategic complementarities to the model by assuming roundabout production of a particular type.[38] We modify firm i's production function to be $Y_{it} = A_{it} N_{it}^{\chi} X_{it}^{1-\chi}$, where X_{it} is the amount of an intermediate input used by firm i. The intermediate input itself is produced from the same CES (36) 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 competitor's prices.

In appendix E.4, we show the following.

Proposition 5. *In the strategic complementarity model, proposition 4 continues to apply unchanged, except that the generalized Phillips curve (46) is now replaced by*

$$\boldsymbol{\pi} = \chi\left(\varphi + \sigma\right)\mathbf{K} \cdot \hat{\mathbf{Y}} \tag{48}$$

[38] Alternative forms of strategic complementarities considered in the menu-cost literature include kinked demand curves (Klenow and Willis 2016) and oligopolistic competition (Mongey 2021). It is an open question how these affect the generalized Phillips curve. Under a Calvo assumption, Gopinath and Itskhoki (2011) derive the implications of kinked demand curves for the aggregate Phillips curve, and Wang and Werning (2020) derive the more complex effects of oligopolistic competition.



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

Figure 12: Impulse response to a monetary shock for the 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).

Hence, as χ declines and strategic complementarity increases, the entire generalized Phillips curve is simply scaled by χ. This result generalizes the existing result from the Calvo model that, in the (NK-PC), strategic complementarities scale down the slope parameter κ. By contrast, our result holds for any generalized Phillips curve $\mathbf{K}$, including for any menu cost models. It therefore provides a useful complement to the recent study of Alvarez et al. (2022b), which characterizes the impulse response of prices in models with strategic complementarity with respect to permanent *nominal* marginal cost shocks, for which no simple scaling result comparable to (48) is attainable.

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 (NK-PC) replaced by the generalized Phillips curve (46). 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 replaced by the generalized Phillips curve $\mathbf{K}$ of either the GL or the NS model, as well as their 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 the empirical distribution of price changes. Instead, the data appears to call for a model with a generalized hazard function, as implied by a more general distribution of menu costs (Alvarez et al., 2022a).[39]

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 GPC of the model can be computed directly from the data without the need to resort to model simulation. This makes the empirical distribution of price changes, together with the overall frequency of adjustment, "sufficient statistics" for the first-order relationship between real marginal cost and inflation in all DSGE models.

Allowing for a general distribution of menu costs. As in section 2.1, we continue to 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 $H(\cdot)$. 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

$$x_{it} = \begin{cases} x_t^* & \text{with probability } \Lambda_t(x_{it-1} - \epsilon_{it}) \\ x_{it-1} - \epsilon_{it} & \text{otherwise, with } \epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2) \end{cases} \tag{49}$$

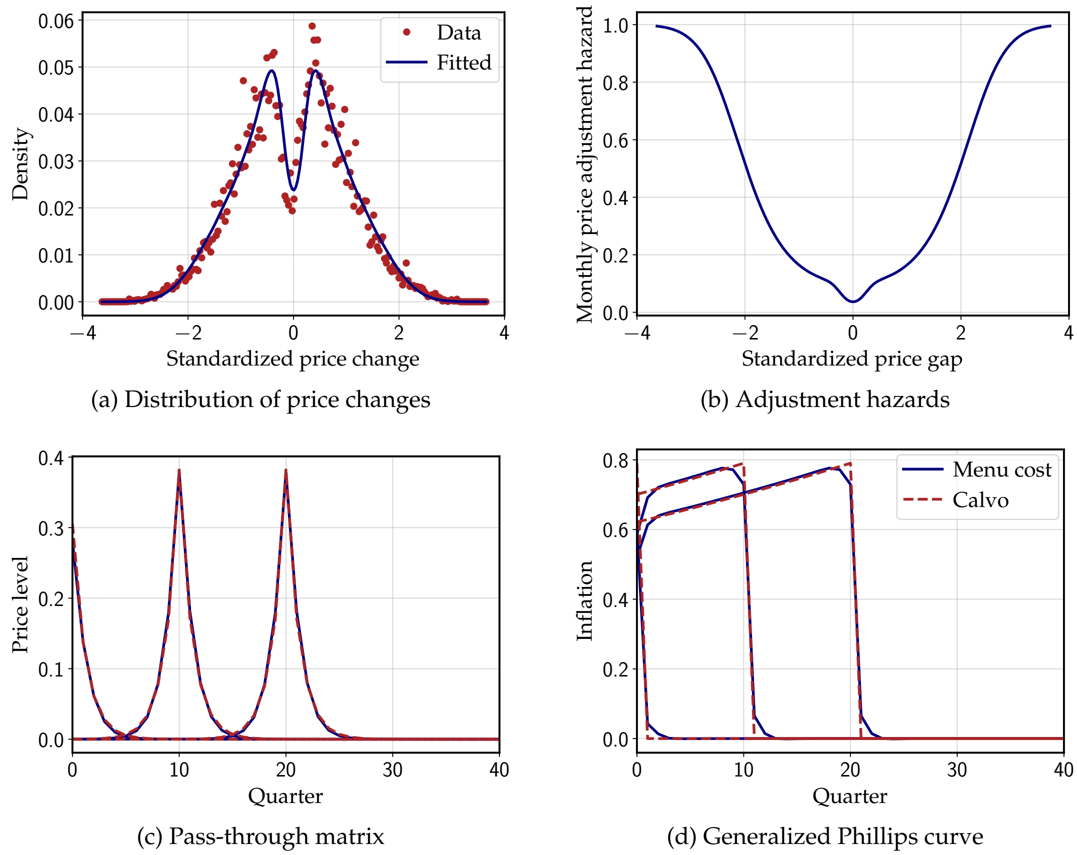
Note that here we assume that ϵ_{it} is drawn from a normal distribution with variance σ_ϵ^2, as in our calibrations if the canonical menu cost model in section 2.5.

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

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



(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.

the following generalization.

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

$$\Psi = freq \cdot \sum_{t=0}^{\infty} E^{t\prime}(0) \cdot \Psi^{\Phi^i} + \int \Lambda'(x) g(x) \cdot \left(\sum_{t=0}^{\infty} E^t(x) \right) \cdot \Psi^{\Phi^e(x)} dx \tag{50}$$

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

Here, the extensive margin in (50) depends on an entire integral of TD pass-through matrices with different survival functions $\Phi_t^e(x)$. However, as we demonstrate next, all objects in (50) can be directly extracted from data on price changes, without any model simulation. In other words, the result in proposition 6 lets us compute the pass-through matrix, and thus also the generalized Phillips curve, based entirely on the distribution as price changes as sufficient statistic.

Backing out generalized hazard and stationary distribution. To back out the objects in (50) from price change data, we proceed as follows. The frequency freq on the right hand side of (50) is readily observable. Conservation of variance implies that $\sigma_\epsilon^2 = \text{freq} \cdot \text{Var}\,(\Delta p)$. Next, we guess a symmetric generalized hazard function $\Lambda(x)$. Given σ_ϵ^2 and $\Lambda(x)$, we compute the stationary distribution $g(x)$ and compare the observed density of price changes Δp with the theoretical density, equal to $\Lambda(-\Delta p)g(-\Delta p)$. We iterate this procedure until we find a suitable generalized hazard function $\Lambda(x)$. Finally, we note that the expected price gap function $E^t(x)$ can be directly computed based on the recovered $\Lambda(x)$ and σ_ϵ^2. This gives us everything needed to evaluate the right hand side of (50), without any model simulation.

Example from Israel. We show how this approach works for data on supermarket prices from Israel. 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 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) recently estimated $\kappa = 0.0031$, and pointed out that typical values in the macro literature are all below 0.05.[40]

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

Simple strategic complementarities, however, cannot solve two broader issues with the New Keynesian Phillips curve: its lack of intrinsic inflation persistence (e.g. Fuhrer and Moore 1995, Galí et al. 2001) and its extreme forward looking property at the heart of the forward guidance puzzle (e.g. Del Negro et al. 2013). Multi-sector models with complex input-output linkages, or

[40] These numbers converts their Table III estimates by assuming that the elasticity of real marginal cost to unemployment is 2, as follows from a calibration of the textbook New Keynesian model in section 5.1 with $\varphi + \sigma = 2$.

deviations from full information rational expectations, 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.

On the theoretical side, our work also leaves open questions. First, comparing the second-order implications of menu cost vs Calvo models would shed light on their relative implications for optimal monetary policy. Finally, the 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).

References

Afrouzi, Hassan and Choongryul Yang, "Dynamic Rational Inattention and the Phillips Curve," *Manuscript*, January 2021.

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*, January 2022, *forthcoming.*

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

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

_ , _ , and Panagiotis Souganidis, "Price Setting with Strategic Complementarities as Mean Field Games," *Manuscript*, March 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.

_ , Katarína Borovičková, and Robert Shimer, "Consistent Evidence on Duration Dependence of Price Changes," Working Paper 29112, National Bureau of Economic Research, July 2021.

_ , 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.

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.

Barro, Robert, "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.

Bonomo, Marco, Carlos Carvalho, Oleksiy Kryvtsov, Sigal Ribon, and Rodolfo Rigato, "Multi-product pricing: Theory and evidence from large retailers in israel," *Manuscript*, 2020.

Caballero, Ricardo J. and Eduardo M. R. A. Engel, "Microeconomic Rigidities and Aggregate Price Dynamics," *European Economic Review*, May 1993, *37* (4), 697–711.

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

_ **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.

Carvalho, Carlos and Felipe Schwartzman, "Selection and monetary non-neutrality in time-dependent pricing models," *Journal of Monetary Economics*, 2015, *76*, 141–156.

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

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," *The Quarterly Journal of Economics*, May 1999, *114* (2), 655–690.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and F. 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.

Fernández-Villaverde, Jesús, "The econometrics of DSGE models," *SERIEs*, 2010, *1* (1), 3–49.

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, David López-Salido, and Nicolas Vincent, "Individual Price Adjustment along the Extensive Margin," *NBER Macroeconomics Annual*, January 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*, 2011, *25* (1), 261–310.

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson, "The Slope of the Phillips Curve: Evidence from U.S. States," *Quarterly Journal of Economics*, February 2022, p. qjac010.

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

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.

La'O, Jennifer and Alireza Tahbaz-Salehi, "Optimal Monetary Policy in Production Networks," Working Paper 27464, National Bureau of Economic Research, July 2020.

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.

Negro, Marco Del, Marc Giannoni, and Christina Patterson, "The Forward Guidance Puzzle," Technical Report, Staff Report, Federal Reserve Bank of New York, May 2013.

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," *Manuscript*, 2020.

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*, 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.

Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics*, April 2003, *50* (3), 665–690.

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.

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," Working Paper 27536, National Bureau of Economic Research, July 2020.

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:

$$dp^*_{it} = -\sigma dW_{it}, \tag{51}$$

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:

$$\rho V(t,x) = \frac{1}{2}(x - MC(t))^2 + \frac{\sigma^2}{2}\partial_{xx} V(t,x) + \lambda \left(V(t,x^*(t)) - V(t,x)\right) + \partial_t V(t,x), \text{ for } x \in (\underline{x}(t), \overline{x}(t)), \tag{52}$$

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:

$$\partial_t g(t, x) = \frac{\sigma^2}{2} \partial_{xx} g(t, x), \tag{53}$$

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

$$p(t) = \int x\, g(t, x)\, dx. \tag{54}$$

By first applying the HJB equation (52), followed by the KFE (53), 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 (51). 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

$$x^*(t) = \frac{1}{\int_0^\infty e^{-\rho s}\Phi(s)\,ds} \int_0^\infty e^{-\rho s}\Phi(s) MC(t+s)\,ds. \tag{55}$$

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

$$p(t) = \frac{1}{\int_0^\infty \Phi(s)\,ds} \int_0^\infty \Phi(s)\, x^*(t-s) ds. \tag{56}$$

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

$$p(t) = \int_0^\infty \Psi(t,s) MC(s)\,ds. \tag{57}$$

We call the operator on the right hand side of (57) the *pass-through operator*. Similarly to the discrete time case, for time-dependent models it is given by the composition of the operators in (55) and (56).

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 (55) and (56), 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 $[0, \infty)$.[41] 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 $(-\bar{x}, \bar{x})$ and the reset point is $x^* = 0$. Let x_t be the price gap of a

[41] 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:

$$\frac{\partial E}{\partial t}(t,x) = \frac{\sigma^2}{2}\frac{\partial^2 E}{\partial x^2}(t,x) - \lambda E(t,x), \tag{58}$$

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

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

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 (58). Moreover, let

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

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

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

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

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

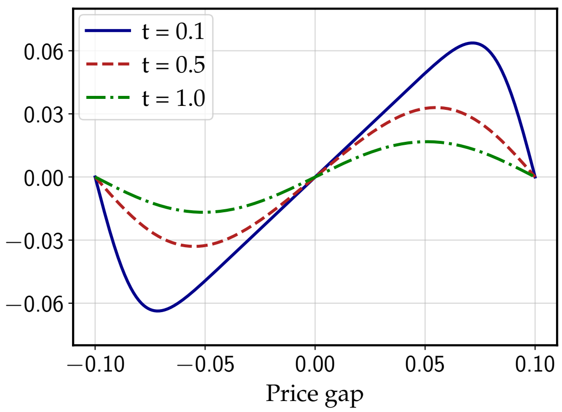
$$\lambda^i(t) = -\frac{\partial}{\partial t} \log \Phi^i(t) = -\frac{\sum_{j \text{ even}} (-1)^{j/2} \eta_j e^{\eta_j t}}{\sum_{j \text{ even}} (-1)^{j/2} e^{\eta_j t}}.$$

From the expressions above, it follows immediately that

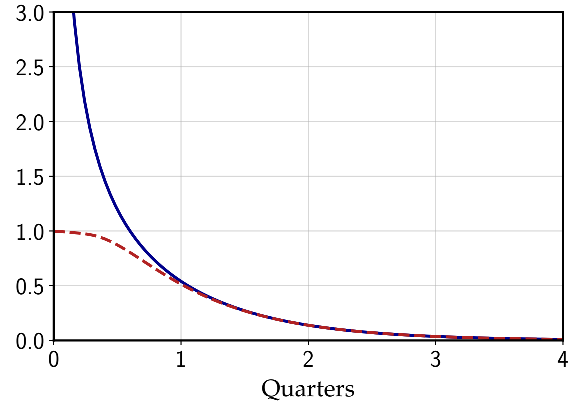
$$\lim_{t \to \infty} \lambda^e(t) = \lim_{t \to \infty} \lambda^i(t) = -\lambda_2,$$

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 $\lim_{t \to 0} \lambda^e(t) = \infty$.[42]

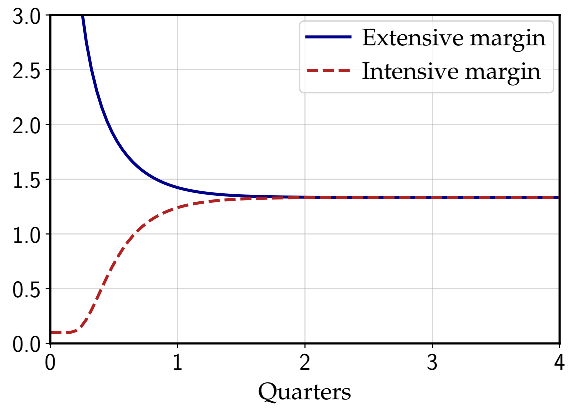
[42] We conjecture that $\lambda^e(t)$ can also be interpreted as the arrival hazard of a point process with countably many points but infinite average density.



(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 $\bar{x} = 0.1$.

B Appendix to Section 2

B.1 Characterizing steady-state policy and distribution

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

$$V^n(x) \equiv \frac{1}{2}x^2 + \beta(1-\lambda)\mathbb{E}\left[\min(V^{n-1}(x+\epsilon), \xi + \min_{x^*} V^{n-1}(x^*))\right] + \beta\lambda \min_{x^*} V^{n-1}(x^*) \qquad (59)$$

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 pricesetter 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 (59) inside $[-M, M]$, or to obtain the optimal policy anywhere. Hence, when analyzing (59), 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 $0 < \bar{x} < M$[43] 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 (59) by the more specific

$$V^n(x) = \frac{1}{2}x^2 + \beta(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x'-x)V^{n-1}(x')dx' \\ + \beta(1-\lambda)(V^{n-1}(0)+\xi)\left(1 - \int_{-\bar{x}}^{\bar{x}} f(x'-x)dx'\right) + \beta\lambda V^{n-1}(0) \qquad (60)$$

where $\bar{x}$ is implicitly determined by the equation $V^{n-1}(\bar{x}) = \xi + V^{n-1}(0)$. It follows directly from (60) and the symmetry of f that V^n will also be nonnegative, symmetric around 0, and satisfy

[43] $\bar{x} < M$ 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 $\xi + V^{n-1}(0)$.

$V^n(x) \geq \frac{1}{2}x^2$ and $V^n(0) \leq \frac{\beta}{1-\beta}(1-\lambda)\xi$. All that remains is investigate the derivative.

Substituting $u = x' - x$, (60) can be rewritten as

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

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

$$\begin{aligned}(V^n)'(x) &= x + \beta(1-\lambda)\left(-f(\bar{x}-x)V^{n-1}(\bar{x}) + f(-\bar{x}-x)V^{n-1}(-\bar{x}) + \int_{-\bar{x}-x}^{\bar{x}-x}(V^{n-1})'(x+u)du\right) \\ &\quad - \beta(1-\lambda)(\xi + V^{n-1}(0))\left(-f(\bar{x}-x)V^{n-1}(\bar{x}) + f(-\bar{x}-x)V^{n-1}(-\bar{x})\right) \\ &= x + \beta(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x'-x)(V^{n-1})'(x')dx' \end{aligned} \tag{61}$$

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 (61) as

$$(V^n)'(x) = x + \beta(1-\lambda)\int_0^{\bar{x}}(f(x'-x) - f(-x'-x))(V^{n-1})'(x')dx' \tag{62}$$

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 (59) 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 (60) holds with this $\bar{x}$.

Now, directly differentiating (60), the continuous differentiability of V follows from that of f. Hence (62) 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 (62) 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 (61) around the fixed point, we get

$$V''(x) = x + \beta(1-\lambda)\int_0^{\bar{x}} (f'(-x'-x) - f'(x'-x))V'(x')dx' \tag{63}$$

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 (63) 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

$$(\mathcal{T}g)(x') = \int_{-\infty}^{\infty} p(x, x')g(x)dx \tag{64}$$

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

$$p(x, x') \equiv \begin{cases} f(x'-x) & |x| \leq \bar{x} \\ f(x') & |x| > \bar{x} \end{cases} \tag{65}$$

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 (64) 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

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

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

$$\|\mathcal{T}(\pi_1 - \pi_2)\| \leq (1 - \mathcal{H})\|\pi_1 - \pi_2\| \tag{66}$$

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 (64)–(65) and the continuous differentiability of f that g is continuously differentiable.

B.2 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

$$\|V\| \equiv \sup_x |V(x)| + \zeta \sup_x |V'(x)| \tag{67}$$

for some $\zeta > 0$. Note that this space is complete (a Banach space).[44] 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[45], given locally by the Bellman equation

$$\begin{aligned} V(x) = \frac{1}{2}(x-c)^2 + \min_{x^*,\underline{x},\bar{x}} \Bigg[& \beta(1-\lambda) \int_{\underline{x}}^{\bar{x}} f(x'-x)V_+(x')dx' \\ & + \beta(1-\lambda)(V_+(x^*) + \zeta) \left(1 - \int_{\underline{x}}^{\bar{x}} f(x'-x)dx'\right) + \beta\lambda V_+(x^*) \Bigg] \end{aligned} \tag{68}$$

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 $d\bar{x} =$

[44] 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's Principles of Mathematical Analysis, Theorem 7.17).

[45] Boundedness of V follows immediately from the Bellman equation, and of of V' follows from differentiating as in (61).

$-dV_+(\bar{x})/V_+'(\bar{x})$, $d\underline{x} = -dV_+(-\bar{x})/V_+'(-\bar{x})$, 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 (68)), where in response to a perturbation dV_+, (68) becomes

$$dV(x) = \beta(1-\lambda)\int_{-\bar{x}}^{\bar{x}} f(x'-x)dV_+(x')dx' + \beta(1-\lambda)\left(1 - \int_{\underline{x}}^{\bar{x}} f(x'-x)dx'\right)dV_+(0) + \beta\lambda dV_+(0) \tag{69}$$

Note that this is a bounded map from dV_+ to dV in our normed space. First, it is immediate from (69) that $\sup_x |dV(x)| \leq \beta \sup_x |dV_+(x)|$.

Second, if we differentiate (69), we obtain

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

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 (67) 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.

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 $d\log MC$, then clearly $\frac{dV(x)}{d\log MC} = -x$ from (68). Then, letting $V_n(x)$ denote the value function with n periods of anticipation of this shock, so that $dV_0(x) = -xd \log MC$, (69) 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

$$dV_n(x) = \beta(1-\lambda) \int_{-\bar{x}}^{\bar{x}} f(x'-x) dV_{n-1}(x') dx' \tag{70}$$

which is equivalent to our envelope result (21).

Observing that (70) 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

$$\frac{dV_n(x)}{dc} = -\beta^n E^n(x) \tag{71}$$

and also

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

Furthermore, the recursion (70) is the same as the recursion (61) obeyed by steady-state $V'(x)$, except that the latter adds x on every iteration (rather than just the base case). Since we have shown that (61) is a contraction with the norm (67), 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 of proposition 2

Define $T : L^2([-\bar{x}, \bar{x}]) \to L^2([-\bar{x}, \bar{x}])$ by

$$(Tg)(x) = \int_{-\bar{x}}^{\bar{x}} f(x'-x) g(x') dx' \tag{72}$$

where the density f is 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). Denote the CDF corresponding to f by $\mathcal{F}$.

Properties of the operator T. We can obtain from (72) several properties of T.

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) 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 property 1. Applying $\psi = \mu^{-1} T\psi$ and (72) again shows that it is C^1.

d) 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 c), 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 $|\lambda| < 1$.

e) 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 $(Tg)(x) = \int_0^{\bar{x}} (\phi(x' - x) - \phi(-x' - x))\, g(x')dx' > 0$, where the inequality follows because single-peakedness implies $\phi(x' - x) - \phi(-x' - x) > 0$ for all $x, x' > 0$, and g is strictly positive on some subset of $[0, \bar{x}]$ of positive measure.

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.

Main result: extensive and intensive margin survival curves. Using the above, we project E^0 onto the basis of eigenfunctions

$$E^0 = \sum_n \langle E^0, \psi_n \rangle \psi_n \tag{73}$$

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 (73) gives

$$E^s(x) = \sum_n \langle E^0, \psi_n \rangle \mu_n^s \psi_n(x) \tag{74}$$

which now holds pointwise for any $s > 0$.[46]

[46]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.

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 (74) that

$$\lim_{s\to\infty} \frac{E^s(x)}{\bar{\mu}^s} = \bar{\psi}(x) \tag{75}$$

Since $E^0(x) = x > 0$ for all $x > 0$, repeatedly applying property e) it must also be true that $E^s(x) > 0$ for all $x > 0$. Taking the limit (75), 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 $x > 0$.

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 (75) and the preceding results have the limits

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

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 $1 - \bar{\mu}$.

Bonus result: 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 e), 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[47] 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

[47] See Theorem 19.3 in Deimling (1985).

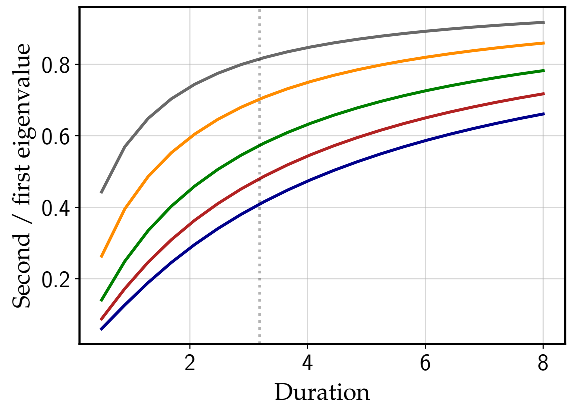


Figure C.1: Ratio of second and first eigenvalues of the transition operator across parameter space of the menu cost model.

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, $1 - \bar{\mu}$.

Application: ratio of second to first eigenvalue.

C.3 Proof of proposition 3

Thanks to our result proving equivalence of an SD model to a mixture of TD models, we can restrict our attention to an arbitrary mixture of (finitely many) TD models. For simplicity, we will start by proving this result for a single TD model, and then show in Step 4 how the argument extends to a mixture.

For a TD model with survival function Φ_s, let us interpret the pass-through matrix Ψ defined in (9), whose columns sum to weakly less than 1, as the transpose of a Markov transition matrix P on the state space of nonnegative integers. In this interpretation, we use the generalized notion of denumerable Markov chain from Kemeny, Snell and Knapp (1976), where the sum of transition probabilities from a state can be less than 1, corresponding to the termination of a chain. (We will use notation and ideas from Kemeny et al. (1976) as well.)

One can rewrite equation (10) as

$$P = AB$$

where

$$A = \frac{1}{\sum_{s\geq 0}\beta^s\Phi_s}\begin{pmatrix} \Phi_0 & 0 & 0 & \cdots \\ \beta\Phi_1 & \Phi_0 & 0 & \cdots \\ \beta^2\Phi_2 & \beta\Phi_1 & \Phi_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and

$$B = \frac{1}{\sum_{s\geq 0}\Phi_s}\begin{pmatrix} \Phi_0 & \Phi_1 & \Phi_2 & \cdots \\ 0 & \Phi_0 & \Phi_1 & \cdots \\ 0 & 0 & \Phi_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The sum $\Psi + \Psi^2 + \ldots$ is then the transpose of the object $\bar{N} \equiv P + P^2 + \ldots$, where $\bar{N}_{ij}$ gives the expected number of visits to state j after being in state i today.

Although $\bar{N}$ is positive, it is not necessarily finite. Our strategy to characterize $\bar{N}$ will be to construct an alternative "shift-invariant" Markov chain with simpler structure and weakly higher transition probabilities, and show that its expected number of visits $\bar{N}^{SI}$ is finite, implying that $\bar{N} \leq \bar{N}^{SI}$ is as well. We will then show that asymptotically, the gap between $\bar{N}$ and $\bar{N}^{SI}$ falls to zero, and so that the asymptotic properties of $\bar{N}^{SI}$ extend to $\bar{N}$.

To do so, let us first extend P, A, and B to cover the state space of all integers, filling in all other transition probabilities with zeros. (This does not alter $\bar{N}_{ij}$ when i and j are both nonnegative.) We then construct the *shift-invariant* Markov chain P^{SI} on the integers[48] with matrix

$$P^{SI} = A^{SI}B^{SI}$$

where A^{SI} and B^{SI} are each defined on all integers such that

$$A^{SI}_{i,j} = \begin{cases} \frac{\beta^{i-j}\Phi_{i-j}}{\sum_{s\geq 0}\beta^s\Phi_s} & i \geq j \\ 0 & i < j \end{cases}$$

and

$$B^{SI}_{i,j} = \begin{cases} 0 & i > j \\ \frac{\Phi_{j-i}}{\sum_{s\geq 0}\beta^s\Phi_s} & i \leq j \end{cases}$$

[48]To interpret (abandoning our Markov chain language for a minute and returning to the original meaning of the pass-through matrix): recall that $P_{s,s+j} = \Psi_{s+j,s}$ gives the amount by which prices will increase at date $s + j$, given a first-order increase in nominal marginal cost at date s that firms learned at date 0. $P^{SI}_{s,s+j}$ is the same, except that all changes in cost are perfectly anticipated going back to date $-\infty$.

Note that both A^{SI} and B^{SI} agree with A and B within the nonnegative integers, but whereas the only nonzero entries in A and B are from the nonnegative integers to themselves, A^{SI} and B^{SI} both have other positive entries. Hence $A^{SI} \geq A$, $B^{SI} \geq B$, and $P^{SI} \geq P$, implying that $\bar{N}^{SI} \geq \bar{N}$ as well.

Also note that both A^{SI} and B^{SI}, and therefore P^{SI}, are *shift-invariant*: all entries along a diagonal are the same, and the probability $P^{SI}_{i,i+j}$ of moving to the right by j only depends on j. Denote this common probability by ψ_j. Similarly, define $a_j \equiv A_{i,i-j}$ and $b_j \equiv B_{i,i+j}$.

Step 1: characterizing z-transform $\psi(z)$ of P^{SI}. Defining the z-transforms $a(z) \equiv \sum_{j=0}^{\infty} a_j z^{-j}$, $b(z) \equiv \sum_{j=0}^{\infty} b_j z^j$, $\psi(z) \equiv \sum_{j=-\infty}^{\infty} \psi_j z^j$, and $\Phi(z) \equiv \sum_{j=0}^{\infty} \Phi_j z^j$, we can write

$$\psi(z) = a(z)b(z) = \frac{\Phi(\beta z^{-1})\Phi(z)}{\Phi(\beta)\Phi(1)} \tag{76}$$

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 12), $\Phi(z)$ 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[49], $|\psi(z)| < \psi(|z|) \leq 1$.

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 2: characterizing $\bar{N}^{SI}$ and its first difference. Since the product of shift-invariant matrices on the integers is given by convolution, and this convolution becomes multiplication when ap-

[49]More explicitly: if $\pi(z) = 1$ and z is complex, then $\pi(z) = |\text{Re}\pi(z)| \leq \sum_{j=-\infty}^{\infty} \pi_j |\text{Re} z^j| < \sum_{j=-\infty}^{\infty} \pi_j |z|^j = \pi(|z|)$, where the final strict inequality holds because we know that, for instance, $\pi_1 > 0$ (which follows from $\Phi_1 > 0$), and for that term $\pi_1 |\text{Re} z| < \pi_1 |z|$ whenever z is not real.

plying the z-transform, it follows that the probability of moving to the right by j after n periods, $[(\bar{P}^{SI})^n]_{s,s+j}$, 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, the expected number of future visits $[\bar{N}^{SI}]_{s,s+j} \equiv [\bar{P}^{SI} + (\bar{P}^{SI})^2 + \ldots]_{s,s+j}$, to a state j to the right of the current one, is given by the jth coefficient of the Laurent series of $\bar{n}(z) \equiv \frac{\psi(z)}{1-\psi(z)} = \sum_{j=-\infty}^{\infty} \bar{n}_j z^j$ in this region. It follows that $\bar{N}^{SI}$ is finite, and hence $\bar{N} \leq \bar{N}^{SI}$ is as well.

Suppose that we are interested in the first difference of the entries of $\bar{N}^{SI}$, i.e. $[\bar{N}^{SI}]_{s,s+j} - [\bar{N}^{SI}]_{s,s+j-1}$. This equals $\bar{n}_j - \bar{n}_{j-1}$, which will be the jth coefficient of the Laurent series

$$k(z) \equiv (1-z)\bar{n}(z) = (1-z)\frac{\psi(z)}{1-\psi(z)} = \sum_{j=-\infty}^{\infty} k_j z^j$$

Since $\frac{\psi(z)}{1-\psi(z)}$ has a simple pole at $z = 1$ (corresponding to the simple zero of $1 - \pi(z)$), 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 $\bar{n}_j$ 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

$$(\beta - z)k(z) \equiv \sum_{j=-\infty}^{\infty} (\beta k_j - k_{j-1}) z^j \tag{77}$$

This removes the simple pole at $z = \beta$, and hence (77) 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

$$\lim_{j\to-\infty} \sup_{l\geq 0} |\beta^j k_j - \beta^{j-l}k_{j-l}| \leq \lim_{j\to-\infty} \beta^{-1}\gamma^j \frac{M}{1-\beta^{-1}\gamma} = 0 \tag{78}$$

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 3: using this to characterize the generalized Phillips curve K. Above, we have already characterized $\bar{N}^{SI}$ and its first difference (in rows). Our goal is now to prove that asymptotically, $\bar{N}^{SI}$ and $\bar{N}$ coincide, in the sense that for any j,

$$\lim_{i\to\infty} [\bar{N}^{SI} - \bar{N}]_{i,i+j} = 0 \tag{79}$$

To prove this, first we derive an expression for $\bar{N}^{SI} - \bar{N}$, writing

$$\begin{aligned}(I-P)P^{SI} - P(I-P^{SI}) &= P^{SI} - P \\ P^{SI}(I-P^{SI})^{-1} - P(I-P)^{-1} &= (I-P)^{-1}(P^{SI}-P)(I-P^{SI})^{-1} \\ \bar{N}^{SI} - \bar{N} &= N(P^{SI}-P)N^{SI}\end{aligned} \tag{80}$$

where in the last line we use $\bar{N} = P + P^2 + \ldots = P(I-P)^{-1}$, $N \equiv I + \bar{N} = (I-P)^{-1}$, and so on.

Let us first characterize the matrix in the middle on the right of (80), $P^{SI} - P$. We recall that $P^{SI} = A^{SI}B^{SI}$ and $P = AB$, and note that actually also $P = AB^{SI}$, since B^{SI} and B coincide for transition probabilities from the nonnegative integers, and A has zero transition probability to negative integers. Hence $P^{SI} = (A^{SI} - A)B^{SI}$.

We next observe that

$$\sum_j (P^{SI}_{ij} - P_{ij}) = \sum_j (A^{SI}_{ij} - A_{ij}) = \frac{\sum_{r=i+1}^{\infty} \beta^r \Phi_r}{\sum_{r=0}^{\infty} \beta^r \Phi_r} \tag{81}$$

where the first equality follows because B^{SI} is stochastic and preserves row sums, and the second equality follows directly from the definitions of A^{SI} and A. We observe that for some sufficiently large C and all $i \geq 0$, (81) is bounded above by $C\beta^i$.

From our earlier characterization, we know that $\bar{N}^{SI}$ and therefore $N^{SI} = I + \bar{N}^{SI}$ has all entries bounded above by some M. It follows that all entries in the ith row of $(P^{SI} - P)N^{SI}$ are bounded by $MC\beta^i$. Using (80), we conclude that

$$\begin{aligned}\lim_{i\to\infty} [\bar{N}^{SI} - \bar{N}]_{i,i+j} &= \lim_{i\to\infty} [N(P^{SI}-P)N^{SI}]_{i,i+j} \\ &\leq \lim_{i\to\infty} [N^{SI}(P^{SI}-P)N^{SI}]_{i,i+j} \\ &\leq \lim_{i\to\infty} \sum_{k=0}^{\infty} N^{SI}_{i,k} MC\beta^k \\ &= MC \sum_{k=0}^{\infty} \beta^k \lim_{i\to\infty} N^{SI}_{i,k} = 0\end{aligned}$$

which, since $\bar{N}^{SI} \geq \bar{N}$, implies that $\lim_{i\to\infty}[\bar{N}^{SI} - \bar{N}]_{i,i+j} = 0$. It follows that

$$\lim_{i\to\infty}(\bar{N}^{SI}_{i,i+j} - \bar{N}^{SI}_{i,i+j-1}) - (\bar{N}_{i,i+j} - \bar{N}_{i,i+j-1}) = 0$$

as well. Since the generalized Phillips curve $K = (I - L)(\Psi + \Psi^2 + \ldots)$ is the transpose of $\bar{N}_{i,i+j} - \bar{N}_{i,i+j-1}$, it follows that its columns asymptotically approach the same two-sided sequence around the diagonal as in the rows of $\bar{N}^{SI}_{i,i+j} - \bar{N}^{SI}_{i,i+j-1}$, which we already characterized in the previous step as the sequence $\{k_j\}$.

Step 4: 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 $\Psi = c_1\Psi^1 + \ldots + c_n\Psi^n$.

Like before, let us interpret the transpose of each pass-through matrix as a Markov transition matrix P^ℓ, and then let us define $P = c_1P^1 + \ldots + c_nP^n$, which is the transpose of Ψ. We will now go through all steps of the previous proof with this P.[50]

First, we can construct $P^{SI,\ell}$ as before for each ℓ, and combine to obtain $P^{SI} = c_1P^{SI,1} + \ldots + c_nP^{SI,n}$, which is still shift-invariant. We then obtain the same characterization of the z-transform $\psi(z)$ of P^{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 1 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 2 is unchanged.

For step 3, the identity (80) remains unchanged, and we can use the argument from (81) to show that for each ℓ, $P^{SI,\ell} - P^\ell$ is bounded above by $C^\ell\beta^i$ for some constant C^ℓ. Taking $C \equiv \max_\ell C^\ell$, it follows that $P^{SI} - P$ is bounded above by $C\beta^i$, and the rest of the proof goes through as before, concluding our argument.

C.4 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

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

[50] Note that P corresponds to a Markov chain where steps are taken according to a random draw from $P^1, \ldots, P^n$, according to the weights $c_1, \ldots, c_n$.

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

$$\phi + f(\overline{x} - x)\overline{x} - f(\underline{x} - x)\underline{x} = F(\overline{x} - x) - F(\underline{x} - x) \tag{82}$$

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

$$f(\overline{x} - x) - f'(\overline{x} - x)\overline{x} = f(\overline{x} + x) - f'(\overline{x} + x)\overline{x} \tag{83}$$

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 (83) holds with $\overline{x} = 1$, then $f(\epsilon) \equiv \hat{f}(\epsilon/\overline{x})$ satisfies (83) with any other $\overline{x} > 0$. With $\overline{x} = 1$, (83) is

$$f(1 - x) - f'(1 - x) = f(1 + x) - f'(1 + x) \tag{84}$$

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 (84). 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 (84) can be rewritten as

$$g'(1 - x) = e^{2x} g'(1 + x) \tag{85}$$

Any positive differentiable g, defined on $[0, 2]$ that satisfies (85) 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 (85) it has to be that

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

Integrating g', we therefore find

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

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

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

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 (82)

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

where we simply used (82) with $x = 0$ (any value of x works).

D Appendix to Section 4

D.1 NK-PC 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:

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \gamma \pi_{t-1} + \kappa \hat{mc}_t + u_t \tag{86}$$

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 (13), 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 86 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.[51] For each SD model, we analyze four different specifications of (86). First, we fix β and κ at the values suggested by approximating the generalized Phillips curve $\mathbf{K}$, 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.

[51]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.

	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.

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 Robustness of the numerical equivalence result

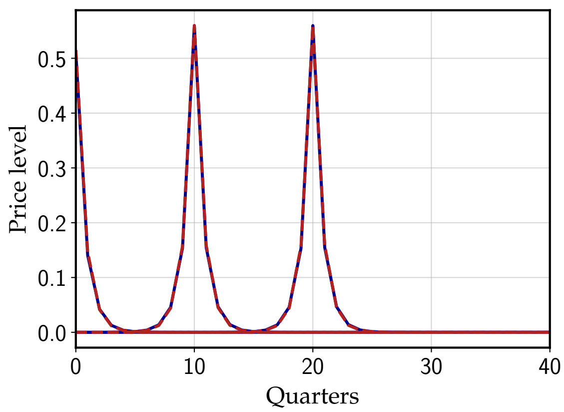
This appendix provides robustness exercises for the numerical equivalence result from section 4. We show that the approximate numerical equivalence holds for several extensions of the baseline menu cost model used in the main text.

First, we introduce steady state 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:

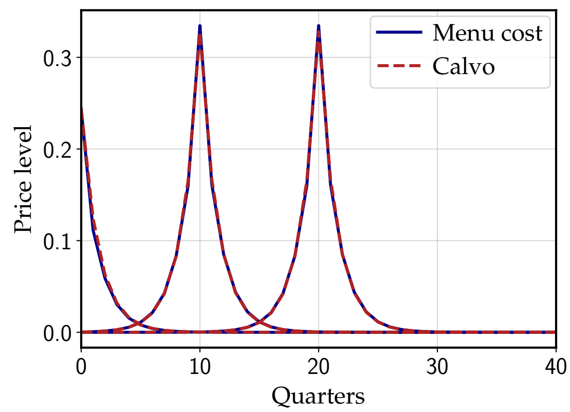
$$p^*_{it} = p^*_{it-1} + \mu + \epsilon_{it}.$$

Since the time unit is one quarter, the drift μ corresponds to an annual inflation rate of 4μ. Figure D.1 shows pass-through and Phillips curve matrices for both Golosov-Lucas and Nakamura-Steinsson 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.

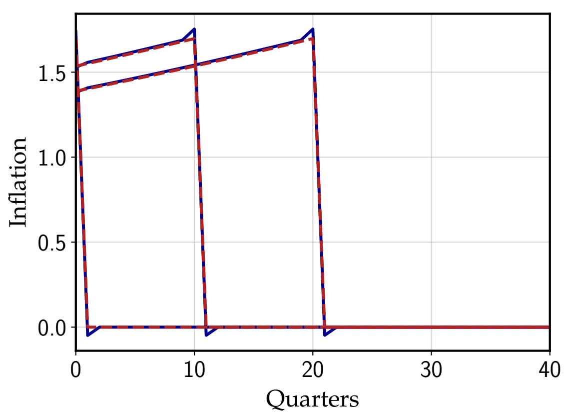
Finally, Figure D.3 shows the best-fitting κ when we calibrate the baseline GL and NS models, and then vary the level of trend inflation μ. The figures show that the slope modestly increases with trend inflation, consistent with the fact that trend inflation increases the steady-state fre-



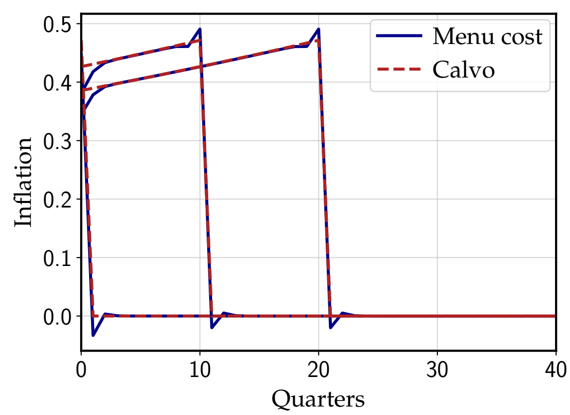
(a) Golosov-Lucas pass-through matrix



(b) Nakamura-Steinsson pass-through matrix



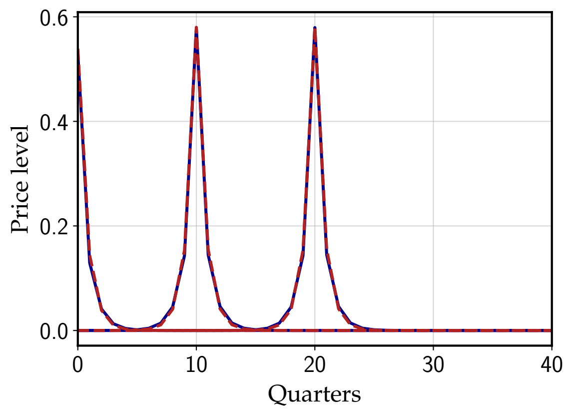
(c) Golosov-Lucas generalized Phillips curve



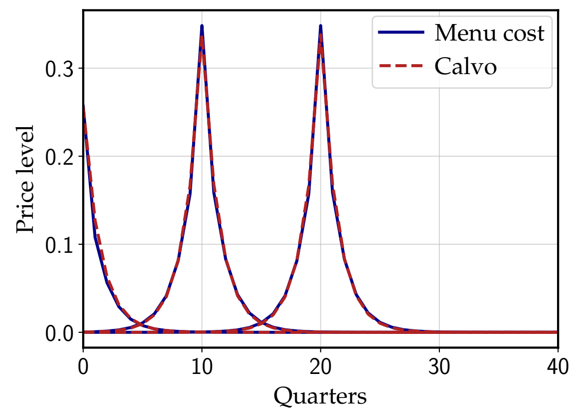
(d) Nakamura-Steinsson generalized Phillips curve

Figure D.1: Menu cost models and Calvo approximations with 2% annual steady state inflation.

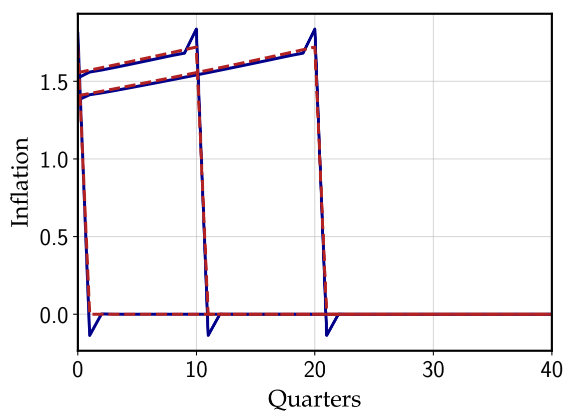
Note: columns $s \in \{0, 10, 20\}$ of the pass-through and Phillips curve matrices for the GL and NS models, calibrated to match the same empirical moments as in the main text, as well as the best-fitting Calvo approximations.



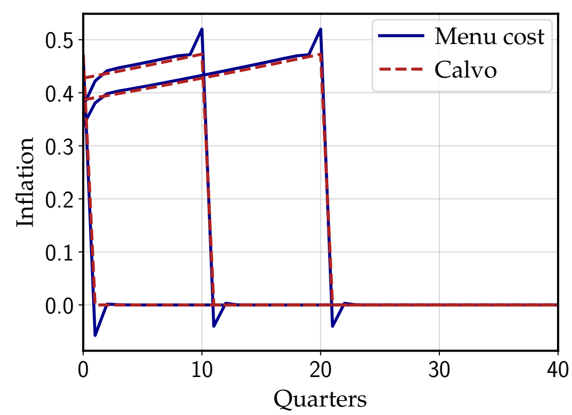
(a) Golosov-Lucas pass-through matrix



(b) Nakamura-Steinsson pass-through matrix



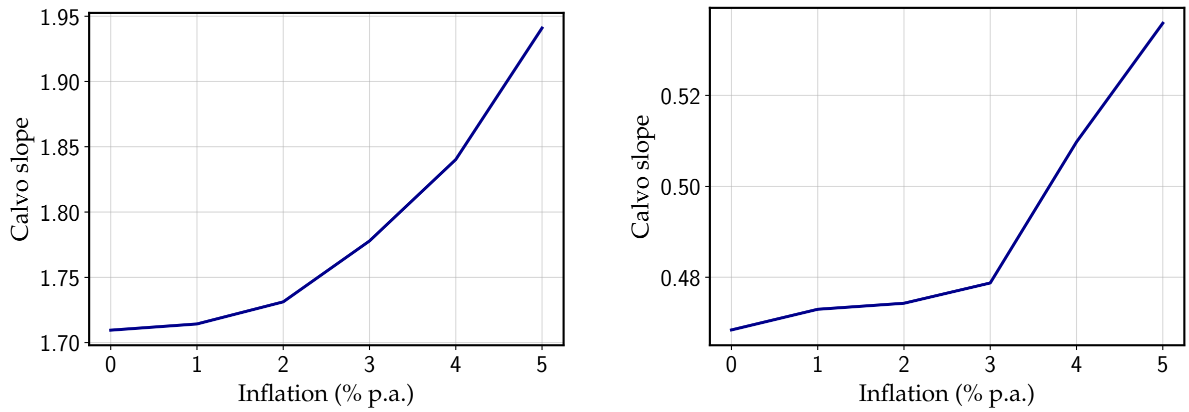
(c) Golosov-Lucas generalized Phillips curve



(d) Nakamura-Steinsson generalized Phillips curve

Figure D.2: Menu cost models and Calvo approximations with 5% annual steady state 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 model

(b) Nakamura-Steinsson model

Figure D.3: Best-fitting Calvo Phillips curve slope for various levels of trend inflation μ

quency of price adjustment (see, e.g., Alvarez, Beraja, Gonzalez-Rozada and Neumeyer 2019).

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

$$\epsilon_{it} = \begin{cases} 0 & \text{with probability } 1-p \\ N(0, \sigma_\epsilon^2) & \text{with probability } p \end{cases}.$$

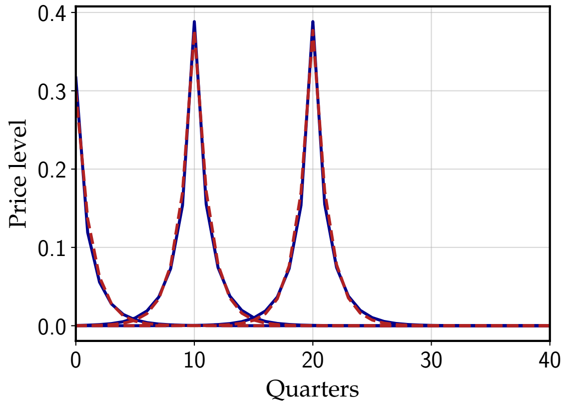
This effectively increases the kurtosis of the (unconditional) shock distribution, so this feature is often referred to in the literature as leptokurtic shocks.[52] Figure D.4 shows results. Similarly to the trend inflation case, the pass-through matrices of the menu cost models are still indistinguishable from their Calvo approximations. The fit of the generalized Phillips curve slightly deteriorates, but is still very good.

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

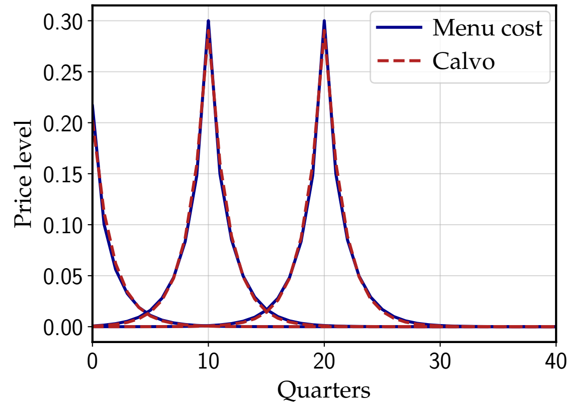
$$\frac{1}{2}\left(x_{it,1} - \log MC_t\right)^2 + \frac{1}{2}\left(x_{it,2} - \log MC_t\right)^2.$$

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.

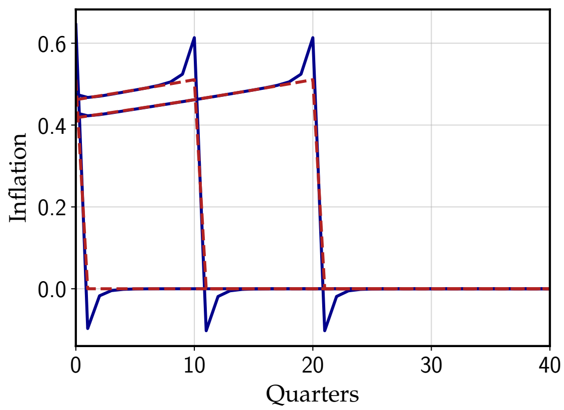
[52] An alternative approach in Karadi and Reiff (2019) assumes a mixture of two normal distributions.



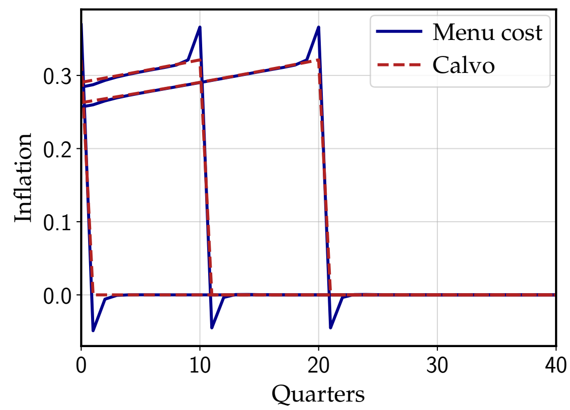
(a) Golosov-Lucas pass-through matrix



(b) Nakamura-Steinsson pass-through matrix



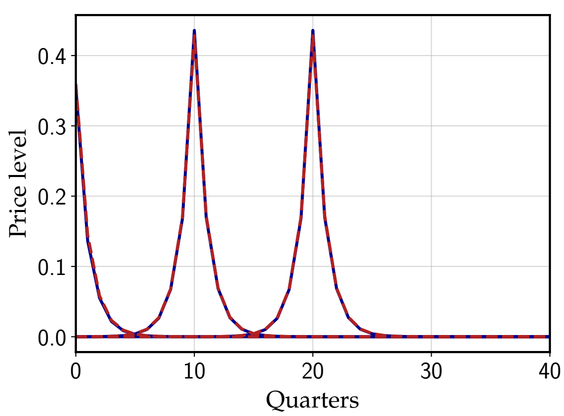
(c) Golosov-Lucas generalized Phillips curve



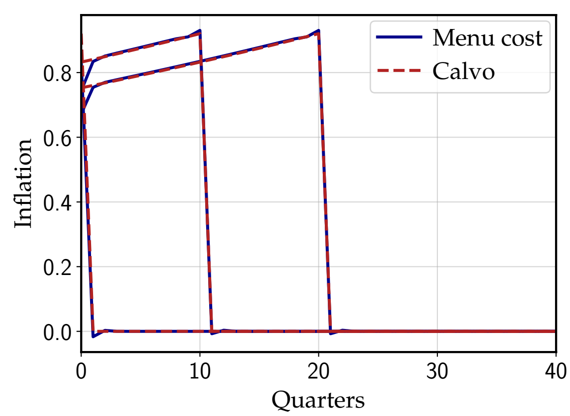
(d) Nakamura-Steinsson generalized Phillips curve

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.



(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.

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 $f_{quarterly} = 1 - (1 - f_{monthly})^3$.

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

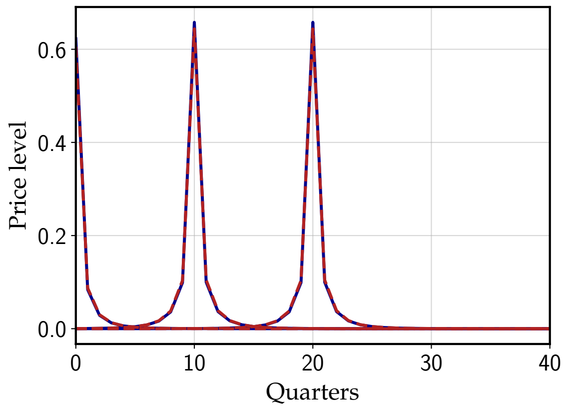
$$p_t = \sum_{j=1}^{N} \omega_j p_{jt},$$

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:

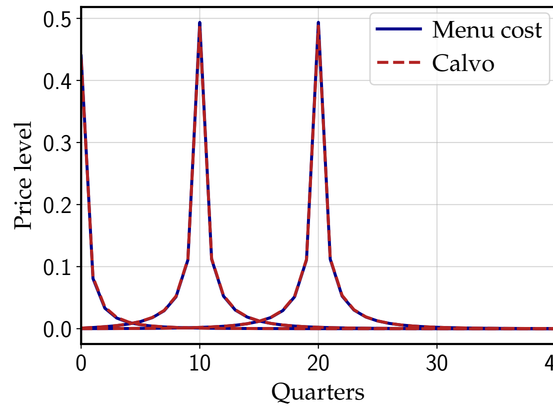
$$\Psi = \sum_{j=0}^{J} \omega_j \Psi_j.$$

Once we have this pass-through matrix, we can apply the same transformation 13 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.

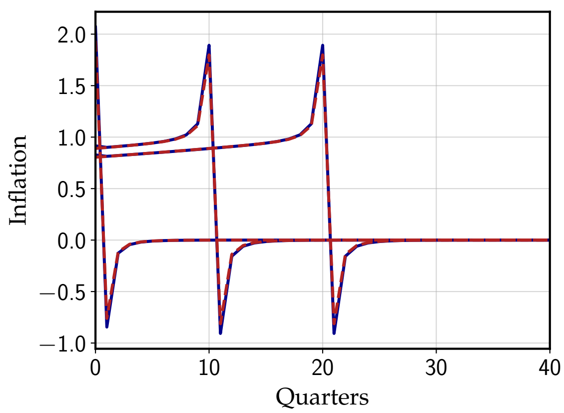
Following the approach outlined above, we calibrate a 14-sector menu cost economy using the same moments as Nakamura and Steinsson (2010), reproduced in table D.2. For each sector, we find the best-fitting Calvo model and compute the corresponding aggregate pass-through and Phillips curve matrices. Figure D.6 shows results. For both our main specifications – with and without free adjustments –, the two models are again almost identical.



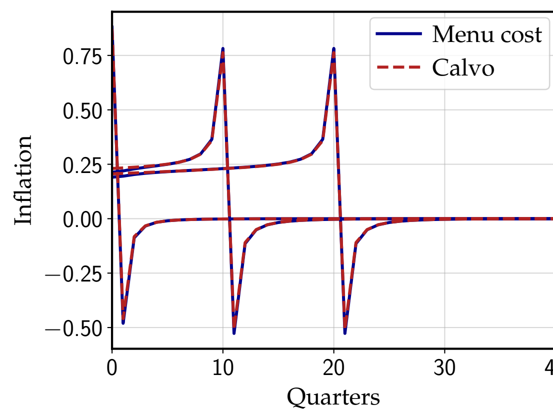
(a) Multi-sector Golosov-Lucas pass-through matrix



(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.

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, 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.[53] 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.

E Appendix to Section 5

E.1 Solving the model with small idiosyncratic shocks

In this section, we derive the quadratic approximation to the objective function (42), and argue that the aggregate menu cost Ξ_t and the price dispersion term Δ_t are second order in σ_ϵ.

E.1.1 Deriving the quadratic approximation

We substitute the deviations of aggregate variables into (42), and denote $\tilde{x}_{it} \equiv \frac{1}{\sigma_\epsilon}(x_{it} - \log W_{ss})$. Then, $\tilde{x}_{it}$ solves the problem:

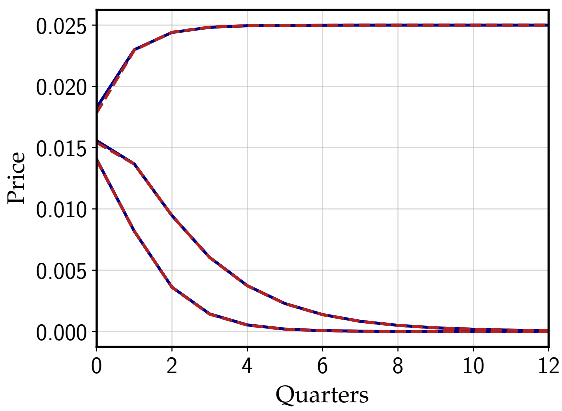
$$\min_{\{\tilde{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(C_{ss} e^{\chi\sigma_\epsilon \hat{C}_t}\right)^{-\sigma} \left[Z_{ss} e^{\chi\sigma_\epsilon \hat{Z}_t} \cdot F\left(\sigma_\epsilon \tilde{x}_{it} - \chi\sigma_\epsilon \hat{W}_t\right) + \sigma_\epsilon^2 \xi_{it} \frac{W_{ss}}{P_{ss}} e^{\chi\sigma_\epsilon \widehat{\frac{W_t}{P_t}}} 1_{\{\tilde{x}_{it} \neq \tilde{x}_{it-1} - \epsilon_{it}\}}\right] \tag{87}$$

We would like to characterize the leading order contribution to $\tilde{x}_{it}$, which we denote by $\hat{x}_{it}$. That is, we'd like to expand $\tilde{x}_{it}$ in σ_ϵ, $\tilde{x}_{it} = \hat{x}_{it} + \mathcal{O}(\sigma_\epsilon)$. We do so by showing that if the objective (87) is expanded in σ_ϵ to second order, the solution is independent of σ_ϵ. This will give us the leading order contribution $\hat{x}_{it}$ to $\tilde{x}_{it}$.

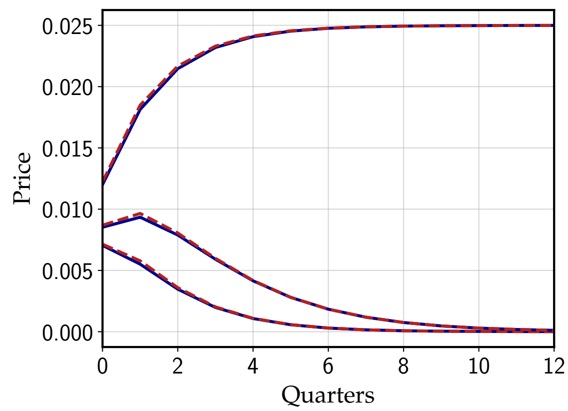
To expand (87) to second order in σ_ϵ, we note that $F'(0) = 0$ and $F''(0) < 0$. Thus,

$$F\left(\sigma_\epsilon \tilde{x}_{it} - \chi\sigma_\epsilon \hat{W}_t\right) = F(0) + \frac{1}{2} F''(0) \cdot \sigma_\epsilon^2 \cdot \left(\tilde{x}_{it} - \chi \hat{W}_t\right)^2 + \mathcal{O}(\sigma_\epsilon^3) \tag{88}$$

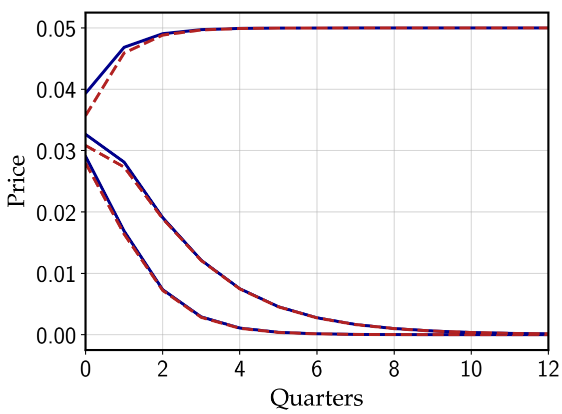
[53] For simplicity, we study only shocks to nominal marginal costs, as shocks to real marginal costs require solving a fixed point problem.



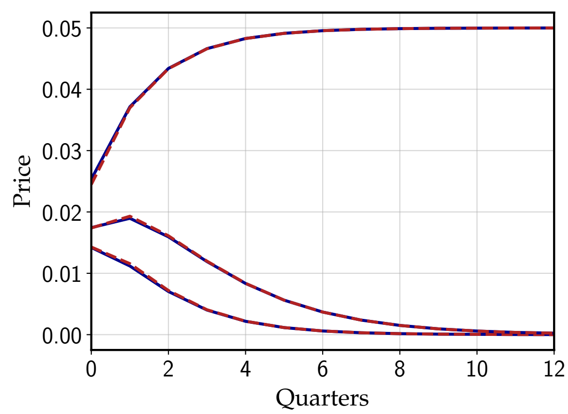
(a) Golosov-Lucas, 2.5% nominal shocks.



(b) Nakamura-Steinsson, 2.5% nominal shocks.



(c) Golosov-Lucas, 5% nominal shocks.



(d) Nakamura-Steinsson, 5% nominal shocks.

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 multi-sector GL and NS models, calibrated as in table 1, as well as linear responses for the best-fitting Calvo approximations. Shock persistence values are $\{0.3, 0.6, 1\}$.

Substituting (88) into (87), we obtain, up to an additive constant

$$\min_{\{\tilde{x}_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t C_{ss}^{-\sigma} \left[Z_{ss} \cdot \frac{1}{2} F''(0) \cdot \left(\tilde{x}_{it} - \chi \hat{W}_t\right)^2 + \xi_{it} \frac{W_{ss}}{P_{ss}} 1_{\{\tilde{x}_{it} = \tilde{x}_{it-1} - \epsilon_{it}\}} \right] \sigma_\epsilon^2 \tag{89}$$

The reason $\hat{C}_t, \hat{Z}_t$, and $\widehat{\frac{W_t}{P_t}}$ do not appear (89) is that they multiply terms that are either constant (and hence irrelevant for the choice of $\tilde{x}_{it}$) or already second order in σ_ϵ^2, in which case they only show up to third order or higher. We denote the solution to (89) by $\hat{x}_{it}$. That solution is independent of σ_ϵ, and is thus the zero-th order term in the expansion of $\tilde{x}_{it}$. If we introduced higher order terms in (89), those would show up as higher-order terms.

Taking these steps together, we found that the expansion of x_{it} in σ_ϵ is given by

$$x_{it} = \log W_{ss} + \sigma_\epsilon \hat{x}_{it} + \mathcal{O}(\sigma_\epsilon^2) \tag{90}$$

E.1.2 Expanding Ξ_t and Δ_t

Observe that the integral $\int_0^1 \xi_{it} 1_{\{x_{it} \neq x_{it-1} - \sigma_\epsilon \epsilon_{it}\}} di$ remains bounded as $\sigma_\epsilon \to 0$, as it converges to $\int_0^1 \xi_{it} 1_{\{\hat{x}_{it} \neq \hat{x}_{it-1} - \epsilon_{it}\}} di$ with $\hat{x}_{it}$ independent of σ_ϵ, given by the solution to (89). Given the formula for Ξ_t, (43), this implies that Ξ_t is at least second order in σ_ϵ.

The formula for Δ_t is given by

$$\Delta_t \equiv \left(\int_0^1 e^{(1-\zeta) x_{it}} di \right)^{\frac{\zeta}{1-\zeta}} \int_0^1 e^{-\zeta x_{it}} di$$

We substitute (90) into the formula to obtain

$$\log \Delta_t = \frac{\zeta}{1-\zeta} \log \left(W_{ss}^{1-\zeta} \int_0^1 e^{(1-\zeta)\sigma_\epsilon \hat{x}_{it}} di \right) + \log \left(W_{ss}^{-\zeta} \int_0^1 e^{-\zeta \sigma_\epsilon \hat{x}_{it}} di \right) + \mathcal{O}(\sigma_\epsilon^2)$$

Simplifying the expression, we find that $\log \Delta_t$ is in fact zero to first order,

$$\log \Delta_t = \zeta \log W_{ss} + \zeta \sigma_\epsilon \int_0^1 \hat{x}_{it} di - \zeta \log W_{ss} - \zeta \sigma_\epsilon \int_0^1 \hat{x}_{it} di + \mathcal{O}(\sigma_\epsilon^2) = 0 + \mathcal{O}(\sigma_\epsilon^2)$$

Thus, any deviation of Δ_t from 1 is second order in σ_ϵ.

E.2 Solving the model with large idiosyncratic shocks

Steady state. We denote steady state variables with subscript "ss". The steady state interest rate is given by

$$\beta (1 + i_{ss}) = 1$$

The steady state distribution $g(x)$ of price gaps is the solution to

$$\min_{\{x_{it}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[Z_{ss} \cdot F\left(x_{it}\right) + \frac{W_{ss}}{P_{ss}} \xi_{it} 1_{\{x_{it} \neq x_{it-1} - \epsilon_{it}\}} \right] \tag{91}$$

where

$$Z_{ss} \equiv \left(\frac{\zeta}{\zeta - 1} \frac{W_{ss}}{P_{ss}} \right)^{1-\zeta} Y_{ss} \tag{92}$$

The real wage is

$$\frac{W_{ss}}{P_{ss}} = \frac{\zeta - 1}{\zeta} \left(\int_0^1 e^{(1-\zeta)x_{it}} di \right)^{-\frac{1}{1-\zeta}} \tag{93}$$

Labor demand is

$$N_{ss} \equiv \left(\frac{\zeta - 1}{\zeta} \right)^{\zeta} Y_{ss} \left(\frac{P_{ss}}{W_{ss}} \right)^{\zeta} \int_0^1 e^{-\zeta x_{it}} di + \Xi_{ss} \tag{94}$$

where aggregate menu cost are given by

$$\Xi_{ss} \equiv (1 - \lambda)\, \xi \left(\int^{\underline{x}} g(x) dx + \int_{\overline{x}} g(x) dx \right) \tag{95}$$

Finally, labor supply satisfies

$$b N_{ss}^{\varphi} = \frac{W_{ss}}{P_{ss}} Y_{ss}^{-\sigma} \tag{96}$$

We normalize $Y_{ss} = 1$. Equations (91)–(96) are then six equations for six unknowns: $g(x)$, Z_{ss}, $\frac{W_{ss}}{P_{ss}}$, N_{ss}, Ξ_{ss}, b.

Dynamics. The equations that characterize the model dynamics are the following. There is a standard Euler equation and Taylor rule as still given by (45) and (47). There is a standard consumption-leisure first order condition:

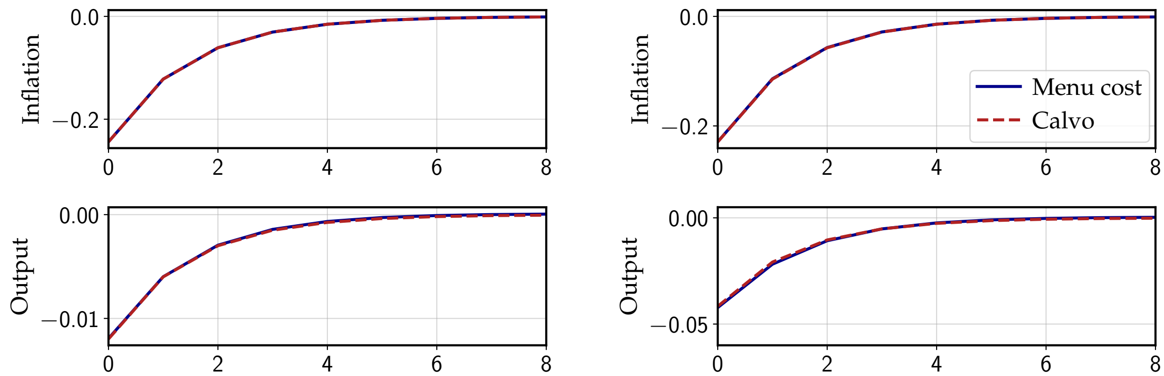
$$\varphi \hat{N}_t + \sigma \hat{Y}_t = \hat{w}_t,$$

where $\hat{w}_t$ is the log real wage relative to steady state. The labor market clearing condition (44) becomes

$$N_{ss} \hat{N}_t \equiv Y_{ss} \Delta_{ss} \hat{Y}_t + Y_{ss} \Delta_{ss} \hat{\Delta}_t + \Xi_{ss} \hat{\Xi}_t.$$

We have so far four equations, but seven variables: $\hat{Y}_t$, $\hat{N}_t$, $\hat{w}_t$, i_t, π_t, $\hat{\Delta}_t$, $\hat{\Xi}_t$. The missing three pieces are the dynamics of π_t, $\hat{\Delta}_t$, and $\hat{\Xi}_t$, which arise from the firms' optimization problem.

From (42), we see that firms now respond not only to shocks to their price gaps, captured by the log W_t term, but also to the shifter Z_t and to real wages, since price adjustment costs are in terms of labor. Using the methodology of Auclert et al. (2021), we are able to represent these responses in terms of infinite Jacobian matrices analogous to (46). Inflation, for instance, is given



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

Figure E.1: Impulse responses for with large idiosyncratic shocks.

Note: both pricing blocks were calibrated to match the same moments as in the main text. The approximating Calvo models are constructed as in section (4).

by

$$\pi = J_{\pi,Y}\hat{\mathbf{Y}} + J_{\pi,w}\hat{\mathbf{w}}, \tag{97}$$

where $J_{\pi,Y}$ captures the response of inflation to changes in output, which appear in the Z_t term, and $J_{\pi,w}$ characterizes the response of inflation to real wage shocks, which affect both price gaps and menu costs. The quantities $\hat{\Delta}_t$ and $\hat{\Xi}_t$ have analogous expressions.

It is interesting to understand why (97) boils down to to (46) when the loss function is quadratic. The reason is that shocks to the value of menu costs or to the discount factor Z_t have no first order effects on inflation. A lower menu cost, for example, generates incentives for firms to adjust their prices now instead of in the future. In the symmetric model, these extra positive and negative adjustments exactly offset each other, and inflation does not respond.

Finally, figure E.1 shows impulse responses of output and inflation for the full general equilibrium model, as well as the best-fitting Calvo approximation. As we can see, results are very close to what we previously had. With large idiosyncratic shocks, the menu cost differs from the Calvo model in two additional dimensions. First, because of the price dispersion term Δ_t. Second, because menu costs are not vanishingly small, and therefore affect total labor demand. These effects, however, are not important quantitatively.

E.3 Proof of proposition 4

TO BE ADDED.

E.4 Proof of proposition 5

Here, we show how to adapt the pass-through matrix and the generalized Phillips curve to account for strategic complementarities. In particular, we prove equation (48). Assume firms have production function

$$Y_{it} = A_{it} N_{it}^{\chi} X_{it}^{1-\chi},$$

where X_{it} is a bundle of intermediate inputs, aggregated as in (36).

Standard cost minimization implies that the aggregate component of firms' nominal marginal costs, already expressed in terms of log deviations from steady state, is

$$\widehat{MC}_t = \chi \widehat{W}_t + (1-\chi)\widehat{P}_t.$$

The parameter χ therefore measures the degree of strategic complementarity—how much of firms' marginal costs depend on other firms' prices—with stronger complementarities corresponding to lower χ.

The price level response to changes in nominal wages is now given by the following fixed point relationship:

$$\begin{aligned} \hat{\mathbf{P}} &= \Psi\left(\chi \widehat{\mathbf{W}} + (1-\chi)\widehat{\mathbf{P}}\right) \\ &= \chi\left(\mathbf{I} - (1-\chi)\Psi\right)^{-1} \Psi \cdot \widehat{\mathbf{W}}. \end{aligned}$$

Therefore, adding strategic complementarities is equivalent to replacing the pass-through matrix by Ψ^{strat}, defined as

$$\Psi^{strat} = \chi\left(\mathbf{I} - (1-\chi)\Psi\right)^{-1} \Psi. \tag{98}$$

The generalized Phillips curve is still obtained from equation (13):

$$\mathbf{K}^{strat} = (\mathbf{I} - \mathbf{L})\,\Psi^{strat}\left(\mathbf{I} - \Psi^{strat}\right)^{-1},$$

By substituting in equation (98), this can be simplified as

$$\begin{aligned} \mathbf{K}^{strat} &= \chi\,(\mathbf{I} - \mathbf{L})\,\Psi\,(\mathbf{I} - \Psi)^{-1} \\ &= \chi \mathbf{K}. \end{aligned}$$

Adding strategic complementarities is thus equivalent to rescaling the generalized Phillips curve by χ, from which (48) follows.

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

$$\Lambda_t(x) = P\left(\xi_{it} \leq V_t(x) - V_t(x_t^*)\right) = \mathcal{H}\left(V_t(x) - V_t(x_t^*)\right). \tag{99}$$

Now expected price gaps evolve according to

$$E^{t+1}(x) = \int_{-\infty}^{\infty} (1 - \Lambda(x')) f(x' - x) E^t(x') dx'$$

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 equation (**??**), we have

$$dP_t^e = -\sum_{s=0}^{t} \int_{-\infty}^{\infty} [d\Lambda_{t-s}(x) g(x) E^s(x)] \, dx. \tag{100}$$

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 equation (**??**).

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 (**??**). Now differentiate (99) 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 (99), 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 $V'(x)$ and $dV_t(x)$ from equations **??** and **??**, respectively. Using the definition $\Phi_t^e(x) = E^t(x)/\overline{x}$, we have

$$-\frac{d\Lambda_t(x)}{\Lambda'(x)} = \frac{\sum_{s=t}^{\infty} \beta^{s-t} \Phi_{t-s}^e(x) d\log MC_s}{\sum_{s=t}^{\infty} \beta^{s-t} \Phi_{t-s}^e(x)}. \tag{101}$$

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 (8).

Now rewrite (100) 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 (9). 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 measurement

In this appendix, we describe the dataset used in section 6, as well as the numerical procedure for recovering the adjustment hazards $\Lambda(x)$. We use the same data as Bonomo et al. (2020). 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 pose any problems for us, as the pass-through matrix is scale-invariant.

In order to recover the adjustment hazards $\Lambda(x)$ from the data, we first need to specify a functional form for it. We postulate

$$\log\left(\frac{\Lambda(x)}{1-\Lambda(x)}\right) = p(x) - s\phi_\sigma(x). \tag{102}$$

In the expression above, $p(x)$ is a 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. Given $\Lambda(x)$, one can compute the stationary distribution of price gaps $f(x)$, which can then be used to compute the resulting distribution of price changes. We then choose the coefficients of $p(x)$, along with the parameters s and σ, in order to minimize the sum of squared errors between empirical and model-implied price-change distributions. Having computed $\Lambda(x)$ and $f(x)$, we then follow the steps outlined in section F.1 to compute the pass-through matrix.