Working Paper Series

Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models

WP 12-09R

Carlos Carvalho
PUC-Rio

Felipe Schwartzman
Federal Reserve Bank of Richmond

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models*

Carlos Carvalho
PUC-Rio

Felipe Schwartzman
Federal Reserve Bank of Richmond

July 30, 2014

Working Paper 12-09R

Abstract

For a given frequency of price changes, the real effects of a monetary shock are smaller if adjusting firms are disproportionately likely to have last set their prices before the shock. This type of selection for the age of prices provides a complete characterization of the nature of pricing frictions in time-dependent sticky-price models. In particular: 1) The Taylor (1979) model exhibits maximal selection for older prices, whereas the Calvo (1983) model exhibits no selection, so that real effects are smaller in the former than in the latter; 2) Selection is weaker and real effects of monetary shocks are larger if the hazard function of price adjustment is less strongly increasing; 3) Selection is weaker and real effects are larger if there is sectoral heterogeneity in price stickiness; 4) Selection is weaker and real effects are larger if the durations of price spells are more variable.

JEL classification codes: E10, E30

Keywords: price setting, monetary non-neutrality, general hazard function, selection effect, heterogeneity.

*For comments and suggestions, we thank participants at AEA 2013, LAMES/LACEA 2012, EEA 2012, SED 2012, ESWM 2012, and ESEM 2011. We also thank Andrew Owens for excellent research assistance. Some of the mathematical results presented here were first derived in Carvalho and Schwartzman (2008) – a legacy working paper that we never submitted for publication. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mails: cvianac@econ.puc-rio.br, felipe.schwartzman@rich.frb.org.

1 Introduction

Infrequent price changes at the microeconomic level do not necessarily imply that monetary disturbances have large real macroeconomic effects. For the same frequency of price changes, the real effects of a monetary shock are small if the firms adjusting their prices are also the ones most likely to change prices by a large amount. The importance of this *selection effect* has been well understood at least since Caplin and Spulber (1987). In their model, large price adjustments by a small fraction of firms completely offset monetary shocks and induce money neutrality. This is because in Caplin and Spulber (1987), as in menu-cost models more generally, there is *self-selection*: firms always have the option of incurring a menu cost to adjust their prices, so that adjusting firms are also the ones which would like to adjust their prices by the greatest amount.[1]

In this paper we argue that selection effects do not necessarily hinge on self-selection. In fact, we show that selection is relevant in time-dependent sticky-price models, where the probability of a price change depends only on the time elapsed since the price was last reset. This is because the real effects of a monetary shock differ depending on whether adjusting firms are more or less likely to have prices that pre-date the shock. More fundamentally, we show that in such an economy, for a given average frequency of price changes, the real effects of a monetary shock depend *solely* on this type of selection. In particular, the real effects of nominal shocks are larger if older prices are relatively less likely to be adjusted.

A proper understanding of the fundamental role of selection in determining the real effects of monetary shocks in time-dependent models is important, since such models are prevalent in the sticky-price literature. While originally used for tractability, subsequent literature has shown that time-dependent pricing rules emerge optimally in the presence of information costs as in Caballero (1989), Bonomo and Carvalho (2004), and Reis (2006).

We tie selection to features of the distribution of price spells, some of which have been singled-out in previous literature as important in determining the extent of monetary non-neutrality in time-dependent price-setting models. In particular, we show that:

1) Calvo (1983) pricing implies no selection, as the probability of price changes does not depend on the age of the price. In contrast, Taylor (1979) pricing implies maximum selection, since changing prices are always the ones which have been in place for longest. This explains why, for a given frequency of price changes, Taylor pricing generates a lower degree of monetary non-neutrality than Calvo pricing (Kiley, 2002).

2) If the hazard of price adjustment is increasing, then selection for older prices is relatively

[1] While Caplin and Spulber (1987) do not consider menu costs explicitly, the state-dependent pricing rule that they postulate can be rationalized by the presence of such costs.

stronger, and the real effects of a nominal shock are smaller than under Calvo pricing. This conforms with discussions by Dotsey, King, and Wolman (1997) and Wolman (1999). Moreover, the more increasing the hazard of price adjustment is, the larger selection effects are.

3) Under certain conditions, we can show that cross-sectoral heterogeneity in price stickiness is associated with lower (and possibly negative) selection, as sectors with low frequency of price changes have both a larger proportion of old prices and a lower probability of price changes. This clarifies and generalizes the finding in Carvalho (2006) that heterogeneity in price setting can lead to larger real effects of monetary shocks.

4) When comparing two economies, one in which the distribution of the duration of price spells is a mean preserving spread of the other, selection for older prices is weaker – and the real effects of the shock are larger – in the economy with more variable price spells. In particular, for a commonly used specification for monetary shocks, the mean and the variance of the duration of price spells are sufficient statistics for the real effects of nominal disturbances.[2]

Our framework encompasses a great degree of generality. As in Dotsey, King, and Wolman (1997, section 3), price changes arrive according to a generic function of the time elapsed since the last price adjustment. We are able to analyze the impact of quite general monetary shocks thanks to an equivalence between the real effects of monetary shocks in sticky-price models and in sticky-information models, as in Mankiw and Reis (2002). For most of the paper, we focus on an environment in which the optimal price for a given firm is neither a strategic substitute nor a strategic complement to the prices set by other firms – what we refer to as *strategic neutrality in price setting*. As a robustness check, we investigate the role of selection in settings with strategic complementarity or substitutability through numerical simulations. The results suggest that the relationship that we uncover between selection and real effects of monetary shocks is robust to those strategic interactions.

Our paper is not the first one to identify a role for selection in time-dependent pricing models. Using a recursive formulation in a discrete-time setting, Sheedy (2010) shows that selection for older prices is associated with higher inflation persistence – an issue that we do not examine. He does not, however, examine the implications of selection for the real effects of monetary shocks. Subsequent work by Alvarez, Le Bihan, and Lippi (2014) elaborates on the link that we uncover between time-dependent selection and monetary non-neutrality in a different setting. Another related paper is Yao (2014), who uses numerical examples to show how differences in the distribution of price durations

[2] Without linking it to selection, we first proved this result in Carvalho and Schwartzman (2008). Subsequently, Vavra (2010) and Alvarez et al. (2012) provided alternative proofs of the same result. The result is also complementary to the one in Alvarez, Le Bihan, and Lippi (2014) which, in a different setup, connects the real effects of monetary shocks to a moment of the distribution of price changes rather than moments of the distribution of price spells.

affect the dynamics of the economy in response to shocks. Finally, Vavra (2010) explores the empirical distribution of price durations estimated from micro-data for the U.S. to study monetary non-neutrality.

We proceed as follows: In Section 2 we lay out the model, which is a continuous time, perfect foresight version of the baseline New Keynesian model, with general distribution of price durations. Section 3 introduces our concept of selection in time-dependent price-setting models, and states the key propositions linking selection to the real effect of a monetary shock. Section 4 shows how selection relates to various ways of summarizing the distribution of price spells, that is, it states and discusses results 1) to 4) listed above. In section 5 we present the numerical results for cases allowing for strategic interactions in price setting. The last section concludes.

2 Model

There is a representative household which derives utility from a continuum of differentiated consumption goods aggregated in a Dixit-Stiglitz composite, and supplies a continuum of firm-specific varieties of labor. Labor is hired by monopolistically competitive firms that produce the goods. The household owns these firms, so it receives back whatever profits they generate. Firms hire labor in competitive markets. We assume a cashless economy with a risk-free nominal bond in zero net supply as in Woodford (2003), and abstract from fiscal policy.

In our analysis, we rely on a first-order approximation of the model around a zero inflation steady state. This allows us to resort to the certainty equivalence principle and focus on the dynamic response of the economy to one-time, zero probability shocks in a world of otherwise perfect foresight. We use a continuous-time formulation since it yields tractable closed form solutions, although none of the key results or intuitions rely on the continuous-time assumption. The representative household maximizes:

$$E_t\left[\int_0^{\infty} e^{-\rho t}\left(\frac{C\left(t\right)^{1-\sigma}-1}{1-\sigma}-\int_0^1 \frac{L_j\left(t\right)^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}}dj\right)dt\right]$$

$$s.t.\ \dot{B}\left(t\right)=i\left(t\right)B\left(t\right)+\int_0^1 W_j\left(t\right)L_j\left(t\right)dj-P\left(t\right)C\left(t\right)+T\left(t\right),\ \text{for } t\geq 0,$$

and subject to a no-Ponzi condition. Here ρ is the discount rate, σ is the inverse of the elasticity of intertemporal substitution, ψ is the Frisch elasticity of labor supply, $C\left(t\right)$ is consumption of the composite good, $L_j\left(t\right)$ is the quantity of labor supplied for the production of variety j, $W_j\left(t\right)$ is the nominal wage for labor of variety j, $T\left(t\right)$ are firms' flow profits received by the consumer, $B\left(t\right)$ denotes bond holdings that accrue a nominal interest at rate $i\left(t\right)$, and $P\left(t\right)$ is a price index to be defined below. E_t is the expectations operator with respect to information available at time t. Given the assumption of perfect foresight except for a one-time, zero probability shock, we can ignore the expectations operator for the solution of the household problem.

The composite consumption good is given by:

$$C(t) \equiv \left[\int_0^1 C_j(t)^{\frac{\varepsilon-1}{\varepsilon}} dj\right]^{\frac{\varepsilon}{\varepsilon-1}},$$

where $C_j(t)$ is consumption of the variety of the good produced by firm j. The elasticity of substitution between varieties is $\varepsilon > 1$. Denoting by $P_j(t)$ the price charged by firm j at time t, the corresponding consumption price index is:

$$P(t) = \left[\int_0^1 P_j(t)^{1-\varepsilon} dj\right]^{\frac{1}{1-\varepsilon}}.$$

The first-order conditions for the representative consumer's optimization problem are:

$$\frac{W_j(t)}{P(t)} = C(t)^{\sigma} L_j(t)^{\frac{1}{\psi}}, \tag{1}$$

$$\frac{\dot{C}(t)}{C(t)} = \sigma^{-1}\left[i(t) - \frac{\dot{P}(t)}{P(t)} - \rho\right],$$

$$C_j(t) = C(t)\left(\frac{P_j(t)}{P(t)}\right)^{-\varepsilon}, \ j \in [0,1]. \tag{2}$$

Firms transform labor into output one for one. They sell their products at a nominal price that they only change infrequently. In the meantime, they commit to producing as much as necessary to satisfy the demand for their output given their chosen price. The timing of those occasional price changes depends probabilistically on the time elapsed since the firm's last price change – i.e., price setting is *time dependent*. Particular examples of time-dependent models include Taylor (1979) and Calvo (1983). We follow Section 3 in Dotsey, King, and Wolman (1997), and consider a general time-dependent setting. We denote the probability of a price surviving for a period of length less than s by a generic cumulative distribution function $G(s)$. At this point, the only restrictions we impose are that $G(s)$ depends only on the time elapsed since the price was last reset, but not on the particular date in which it was reset. Note that G being a c.d.f. implies that $\lim_{s\to\infty} G(s) = 1$, so that all price spells come to an end with probability one. Certain results require additional restrictions on G that we will introduce as needed.

A firm that resets its price at time t chooses the price $X_j(t)$ to solve:

$$\max_{X_j(t)} E_t\left[\int_0^\infty e^{-\rho s}(1-G(s))\left[X_j(t)Y_j(t+s)-W_j(t+s)N_j(t+s)\right]ds\right]$$
$$s.t.\ Y_j(t+s)=N_j(t+s), \tag{3}$$
$$Y_j(t+s)=\left(\frac{X_j(t)}{P(t+s)}\right)^{-\varepsilon}Y(t+s),$$

where $N_j(t+s)$ is the amount of labor demanded by the firm, and where the demand function already takes into account that goods market clearing implies $C_j(t)=Y_j(t)$. The first-order condition yields:

$$X_j(t)=E_t\left[\frac{\varepsilon}{\varepsilon-1}\frac{\int_0^\infty e^{-\rho s}(1-G(s))P(s)^{\varepsilon}Y(s)W_j(s)ds}{\int_0^\infty e^{-\rho s}(1-G(s))P(s)^{\varepsilon}Y(s)ds}\right].$$

As is usual in the literature, we focus on the symmetric equilibrium in which all adjusting firms choose the same nominal price. This allows us to drop the j subscripts and denote the price set by any firm at time t as $X(t)$. Moreover, we assume uniform staggering of pricing decisions, so that the price index satisfies:

$$P(t)=\left[\int_{-\infty}^{t}\Lambda(1-G(t-v))X(v)^{1-\varepsilon}dv\right]^{\frac{1}{1-\varepsilon}},$$

where $\Lambda dt\equiv\left[\int_0^\infty(1-G(s))ds\right]^{-1}dt$ is the constant "fraction" of prices which are changed over an infinitesimally small interval dt. We refer to Λ as the average frequency of price changes in the economy. Using integration by parts, it is straightforward to show that:

$$\Lambda^{-1}=\int_0^\infty s dG(s),$$

which is the average duration of price spells.

The model is closed by a monetary policy specification that ensures existence and uniqueness of a rational expectations equilibrium. Following standard practice in the price-setting literature (e.g. Mankiw and Reis, 2002), we leave the details of monetary policy unspecified and assume an exogenous path for nominal aggregate demand, $M(t)=P(t)Y(t)$.

We log-linearize the model around a zero-inflation steady state. In this log-linear environment, firms that change prices at time t set (lowercase variables denote log-deviations from the steady state):

$$x(t)=x_j(t)=E_t\left[\frac{\int_0^\infty e^{-\rho s}(1-G(s))w_j(t+s)ds}{\int_0^\infty e^{-\rho s}(1-G(s))ds}\right]. \tag{4}$$

Log-linearizing the labor supply condition in equation (1), and combining the log-linear versions of the production function (equation 3), the household's demand for varieties (equation 2), and the market clearing condition $C_j(t) = Y_j(t)$ yields the following equilibrium expression for nominal wages:

$$w_j(t+s) = p(t+s) + \left(\sigma + \psi^{-1}\right) y(t+s) - \varepsilon\psi^{-1}(x(t) - p(t+s)).$$

Note that $w_j(t+s)$ is the same for all j, so that, consistent with the symmetry assumption above, $x_j(t)$ is also the same for all j. We can also use $m(t+s) = p(t+s) + y(t+s)$ to substitute out $y(t+s)$, rearrange slightly, and obtain:

$$w_j(t+s) = \left(1 - \sigma - \psi^{-1}\right) p(t+s) + \left(\sigma + \psi^{-1}\right) m(t+s) - \varepsilon\psi^{-1} x(t+s).$$

Substituting the expression above in the first-order condition for the firm's problem (equation 4) and rearranging yields:

$$x(t) = E_t\left[\frac{\int_0^\infty e^{-\rho s}(1 - G(s))\left[\alpha m(t+s) + (1-\alpha) p(t+s)\right] ds}{\int_0^\infty e^{-\rho s}(1 - G(s))\, ds}\right], \tag{5}$$

where $\alpha = \frac{\sigma + \psi^{-1}}{1 + \varepsilon\psi^{-1}}$.

According to equation (5), the model implies *strategic neutrality* in price setting if $\alpha = 1$. This means that the marginal cost of production for a given firm and, therefore, its desired price, only depends on the exogenous process $m(t+s)$ and not on decisions made by other firms. This requires specific constellations of primitive parameters such as, for example, $\sigma = 1$ and $\psi \to \infty$ (log utility in consumption and linear disutility of labor). More generally, pricing decisions will be either strategic substitutes or strategic complements. If $\alpha < 1$, there is strategic complementarity in price setting, meaning that firms will choose prices close to what they expect the aggregate price level to be. With $\alpha > 1$ pricing decisions are strategic substitutes.

Finally, the aggregate price level is given by:

$$p(t) = \int_{-\infty}^{t} \Lambda(1 - G(t-v))\, x(v)\, dv. \tag{6}$$

2.1 Monetary shocks

The economy starts with a constant level of nominal aggregate demand M^{old}, with associated pricing decisions X^{old}, the aggregate price level P^{old}, and constant output Y^{old}. We analyze the impact of a one-time, unforeseen shock to nominal aggregate demand. The shock hits the economy at $t = t_0$, yielding thereafter a new path for nominal aggregate demand $M^{new}(t)$, and associated paths for pricing decisions, aggregate price level, and output – respectively, $X^{new}(t)$, $P^{new}(t)$, and $Y^{new}(t)$. The assumptions that price setting is purely time dependent and that price changes are uniformly staggered over time allow us to set, for notational convenience, $t_0 = 0$ without loss of generality.

In log-linear terms, the ex-post path of nominal income is:

$$m(t) = \begin{cases} m^{old}, \text{ if } t < 0, \\ m^{new}(t), \text{ if } t \geq 0. \end{cases} \tag{7}$$

The assumption of a one-time unforeseen shock implies that $E_t[M(t+s)] = M^{old}$ if $t < 0$ and $E_t[M(t+s)] = M^{new}(t)$ if $t \geq 0$, and analogously for $X(t)$, $P(t)$ and $Y(t)$. Thus, from this point onward, we drop the expectations operator and use the superscripts "new" and "old" instead.

3 Selection and monetary non-neutrality

In this section we introduce a concept of selection appropriate for time-dependent models. We show that, for a given frequency of price changes, this type of selection contains the same information as the distribution of price durations, so that a characterization of how selection evolves over time provides a complete description of nominal rigidities in the model. Finally, we show how selection affects monetary non-neutrality, with lower selection for prices set before the shock being associated with higher real effects of nominal shocks.

3.1 Selection

In statistics, there is a selection bias if a sample is not a random draw from the population. In that case, sample moments provide biased estimates of population moments. By analogy, the prices being reset at a given point in time are a sample of the population encompassing all existing prices. As a measure of selection bias, we focus on the fraction of prices set before the shock ("old prices") being reset at t, as compared to the corresponding fraction of old prices in the population still in place at t.

Because the distribution of the duration of price spells, G, is time-invariant, at any time $t \geq 0$, the fraction of old prices among changing prices is equal to $1 - G(t)$ – which is the probability that a price survives for t or longer. In turn, $1 - \omega(t) \equiv 1 - \int_0^t \Lambda(1 - G(s))\, ds$ is the fraction of old prices in the population at time t. In this context, we say there is *positive* selection for old prices if $1 - G(t) > 1 - \omega(t)$ and *negative* selection otherwise. This suggests a natural measure of selection for old prices at each point in time after a shock.

Definition 1. *For all t such that $\omega(t) < 1$, selection (at t), denoted by $\mu(t)$, is defined as*

$$\mu(t) \equiv \frac{1 - G(t)}{1 - \omega(t)} - 1,$$

and for t such that $\omega(t) = 1$,

$$\mu(t) = 0.$$

The extension of the definition for the cases in which $\omega(t) = 1$ is natural, since with $\omega(t) = 1$ all adjusting prices as well as all prices in the population are set after the shock (that is, they are all "new"). Hence, the "sample" of prices which can adjust at any point in time has the same composition as the population and there is no selection bias.

The state of the economy at any $t \geq 0$ is a function of the history of selection for old prices starting at the time of the shock. To capture this history, we also employ a related measure,which emphasizes not selection at a given point in time, but cumulative selection since the shock hit:

Definition 2. *Cumulative selection (at t), denoted by $\Xi(t)$, is defined as*

$$\Xi(t) \equiv \int_0^t \mu(s)\, ds.$$

We refer loosely to selection for old prices in economy A being stronger than in economy B if either $\mu_A(t) > \mu_B(t)\ \forall t$ and/or $\Xi_A(t) > \Xi_B(t)\ \forall t$. It is easy to see that the first ordering implies the second, but that the converse is not necessarily true.

We now proceed to show how the population of old prices at any point in time is determined by the history of selection up to that point. After the monetary shock hits, the pool of new prices $\omega(t)$ increases as firms *a)* have the opportunity to change prices (this is given by the frequency of price changes, Λ) and *b)* are doing so for the first time after the shock (this applies to a fraction $1 - G(t)$ of price changers). Therefore:

$$\frac{\partial \omega(t)}{\partial t} = \Lambda (1 - G(t)). \tag{8}$$

Solving the differential equation (8) with $\omega(0) = 0$ as a boundary condition, and using the definitions above yields the following:[3]

$$1 - \omega(t) = e^{-\Lambda t - \Lambda \int_0^t \mu(v) dv} = e^{-\Lambda t - \Lambda \Xi(t)}. \tag{9}$$

Equation (9) suggests that, given Λ, G can be obtained from μ (and vice versa). As the following lemma shows, this is indeed the case:[4]

[3]See Lemma A.1 in appendix for a formal statement and proof.
[4]All proofs are in the Appendix.

Lemma 1. *Let $\mathcal{M}$ be the set of all functions $\mu : [0,\infty) \to \mathbb{R}$ that can be constructed using **Definition 1**. Let $\mathcal{G}^{\Lambda}$ be the set of all functions $G : [0,\infty) \to [0,1]$ satisfying $\int_0^{\infty} (1 - G(s))\, ds = \Lambda^{-1}$. Then, there is a mapping $f : \mathcal{M} \to \mathcal{G}^{\Lambda}$ that allows us to recover G from μ.*

The last result implies that, given the frequency of price changes Λ, μ and G are equally valid primitives for the general class of time-dependent pricing models that we consider.

3.2 Monetary non-neutrality

We now examine the link between selection and monetary non-neutrality. We focus on the case with strategic neutrality in price setting, since it allows us to isolate the role of selection from the well-known effects of interactions between firms' pricing decisions. Using numerical simulations, in Section 5 we examine whether our main results survive the presence of pricing interactions.

The effects of the shock on real output are given by:

$$y^{new}(t) - y^{old}. \tag{10}$$

We measure the degree of monetary non-neutrality by the discounted cumulative effect of the shock on output. More specifically, our measure of non-neutrality is given by:

$$\Gamma = \int_0^{\infty} e^{-\rho t} \left[y^{new}(t) - y^{old} \right] dt.$$

In the Appendix we show that, up to a first-order approximation, this measure is proportional to the ex-post utility impact generated by the monetary shock. We refer to Γ generically as *the real effects of the monetary shock.*

3.2.1 Level shocks

We start by analyzing the commonly used case where, following the shock, the level of nominal income changes once and for all, that is, $m^{new}(t) = m^{new} = m^{old} + \Delta m$ for some constant Δm. Apart from being a common benchmark, this case is interesting because the link between selection and the real effects of monetary shocks is particularly transparent.

From (5) it follows immediately that:

$$x^{old} = m^{old},\ x^{new} = m^{new}.$$

Taking into account the different price-setting decisions made before and after $t = 0$, we can then write the evolution of the aggregate price level for $t \geq 0$ as:

$$p^{new}(t) = p(t) = \omega(t)\, m^{new} + (1 - \omega(t))\, m^{old},$$

where $\omega(t)$ is the fraction of firms with new prices in the population (i.e., who last set their prices after the shock).

The effects of the shock on real output are thus given by:

$$y^{new}(t) - y^{old} = m^{new} - p^{new}(t) - \left(m^{old} - p^{old}\right) = \Delta m\,(1 - \omega(t)).$$

In words, the output effect at t is proportional to the size of the shock Δm and to the fraction of firms with old prices at t, $1 - \omega(t)$. Thus, for a given sized shock, the real effects at t are larger if the pool of old prices is larger.

Using (10), we can write the real effects of the shock as:

$$\Gamma = \Delta m \int_0^\infty e^{-\rho t}(1 - \omega(t))\, dt. \tag{11}$$

The real effects are increasing in the integral over time of the fraction of old prices in the population. The longer the fraction of old prices in the population takes to shrink to zero after the shock, the larger are its real effects. It follows from (9) that:

$$\frac{\Gamma}{\Delta m} = \int_0^\infty e^{-\rho t}(1 - \omega(t))\, dt = \int_0^\infty e^{-(\rho+\Lambda)t - \Lambda \int_0^t \mu(v)dv} dv = \int_0^\infty e^{-(\rho+\Lambda)t - \Lambda \Xi(t)} dt. \tag{12}$$

We can thus derive the following immediate implications, summarized in the lemma below:

Lemma 2. *Given Λ and strategic neutrality ($\alpha = 1$), the effects of a shock to the level of nominal income ($m_1 = m_0 + \Delta$) are larger if either*

1) selection, $\mu(t)$, is smaller for all t, or

2) cumulative selection, $\Xi(t)$, is smaller for all t.

3.2.2 General shocks

One difficulty in establishing analytical results for general shocks is that, differently from the simple case with a level shock, the cross-sectional distribution of prices is in general not concentrated on only two values – one for old prices and one for new prices. In spite of that, we are able to handle these cases thanks to the following proposition:[5]

[5] Since we rely on a log-linear approximation to the model around a zero inflation steady state, these more general shocks should not involve permanently non-zero inflation.

Proposition 1. *Consider an economy characterized by a distribution of price spells G and strategic neutrality ($\alpha = 1$). The real impact of a monetary shock of the general form considered in equation (7) is*

$$\Gamma = \int_0^\infty e^{-\rho t} (1 - \omega(t)) \left(m^{new}(t) - m^{old}\right).$$

This proposition holds in spite of the fact that, in general, $y^{new}(t) - y^{old} \neq (1 - \omega(t)) \left(m^{new}(t) - m^{old}\right)$. The fact that it holds is a consequence of optimality of firms' price-setting decisions. Given strategic neutrality in price-setting, a firm j choosing its price after the shock would like to set $x_j(t) = m^{new}(t+s)$ for all s, but this is impossible if $m^{new}(t+s)$ varies over time. As a "compromise", it optimally sets $x_j(t)$ to be equal to a weighted average of $m^{new}(t+s)$, with weights given by the probability with which it expects the price to remain in place at each date $t+s$. For some period of time, $x_j(t)$ will remain below $m^{new}(t+s)$, and for some other period it will remain above. Over time these differences, as weighted by the probabilities of the price remaining in place, cancel out exactly, so that overall the real effects are the same as if the firm was able to set $p_j(t+s) = m^{new}(t+s)$ for all $s \geq 0$.

In the Appendix, we show that Proposition 1 can be alternatively formulated as stating that the real effects of a monetary shock in a sticky-price model are identical to those effects in a sticky-information model, so long as the distribution of price spells in the former is identical to the distribution of price plans in the latter. Thus all the analytical results in this paper translate to an equally large class of models with sticky information, as in Mankiw and Reis (2002).

As in Section 3, we can write the result in Proposition 1 in terms of cumulative selection:

$$\Gamma = \int_0^\infty e^{-(\rho+\Lambda)t - \Lambda\Xi(t)} \left(m^{new}(t) - m^{old}\right) dt. \tag{13}$$

From equation (13), it is evident that, given Λ, so long as $m^{new}(t) > m^{old}$ for all t (or vice versa), the monetary shock has smaller real effects in the economy with larger cumulative selection $\Xi(t)$ everywhere. This last result implies a generalization of Lemma 2:

Proposition 2. *Consider a shock to nominal aggregate demand characterized by $m^{new}(t) \geq m^{old}$ for all t.*

Consider the impact of the shock in two economies, A and B, characterized by distributions of price durations $G_A(t)$ and $G_B(t)$, with $\int_0^\infty (1 - G_A(t))\, dt = \int_0^\infty (1 - G_B(t))\, dt = \Lambda^{-1}$. Then, $\Gamma_A < \Gamma_B$ if either

1) $\mu_A(t) \geq \mu_B(t)$ $\forall t$ or,

2) $\Xi_A(t) \geq \Xi_B(t)$ $\forall t$.

4 Selection and the distribution of price durations

We now turn to results concerning how selection is related to different properties of the distribution of price spells. We start by discussing the benchmark cases of Taylor and Calvo pricing. We then revisit two topics which have been the subject of previous work: the slope of hazard functions and ex-ante heterogeneity in price setting. Finally, we show that there is a link between selection and the variance of price durations, and explore conditions under which this variance is an accurate scalar measure of the real effects of monetary policy shocks.

4.1 Benchmark cases: Taylor and Calvo pricing

We start with a discussion of the two most widely used time-dependent models, which are the ones proposed by Taylor (1979) and Calvo (1983). We show that these cases are polar opposites insofar as selection is concerned. In particular, Taylor pricing implies maximal selection and Calvo pricing implies no selection. Thus, the real effects of monetary shocks will be minimal under Taylor and larger in Calvo than in any model with non-negative selection.

4.1.1 Taylor pricing

Firms set prices for a fixed period of time (given by Λ^{-1}). Thus, the distribution of price durations is degenerate at Λ^{-1}. This specification has been very influential in the sticky-price literature, and, apart from Taylor (1979), it has been used in prominent papers such as Chari, Kehoe, and McGrattan (2000).

For a given frequency of price durations Λ, we can define Taylor pricing in terms of our notation as:

$$G^{Taylor}(t) = \begin{cases} 0 \text{ if } t < \Lambda^{-1}, \\ 1 \text{ otherwise.} \end{cases}$$

Under Taylor pricing, selection at time t is:

$$\mu^{Taylor}(t) = \begin{cases} \frac{1}{1-\Lambda t} - 1 \text{ if } t < \Lambda^{-1}, \\ 0 \text{ otherwise.} \end{cases} \tag{14}$$

Selection is equal to zero for $t \geq \Lambda^{-1}$ since from that point onward the pool of old prices is thoroughly depleted, so that $1 - \omega^{Taylor}(t) = 0$. Selection is positive elsewhere. Within the range where selection is positive, it is also maximal, since all changing prices were set before the shock. We formalize the point in the following Lemma:

Lemma 3. *Consider an arbitrary time-dependent economy with distribution of price durations characterized by $G(t)$ and with average frequency of price durations Λ. Let $\mu(t)$ and $\Xi(t)$ be,* *respectively, the corresponding selection and cumulative selection functions. Let* $\mu^{Taylor}(t)$ *and* $\Xi^{Taylor}(t)$ *be, respectively, the selection and cumulative selection functions for a Taylor economy with average frequency of price changes* Λ. *Then* $\mu^{Taylor}(t) \geq \mu(t)$ *for all* $t < \Lambda^{-1}$ *and* $\Xi^{Taylor}(t) \geq \Xi(t)$ *for all* t.

Given Proposition 2, it follows immediately that, for a given Λ, Taylor pricing implies the smallest real effects among all time-dependent pricing models.[6]

4.1.2 Calvo pricing

A further leading example of time-dependent pricing used in the literature is the one proposed by Calvo (1983), which is the key building block of the canonical New Keynesian model. In this setting, the probability of a given firm changing its price over any given period of time does not depend on the time elapsed since it last adjusted. This implies an exponential decay of the survival probability of a price.

In terms of our notation, we can denote the cumulative distribution of price durations under Calvo as:

$$G^{Calvo}(t) = 1 - e^{-\Lambda t}.$$

It is easy to verify that:

$$\omega^{Calvo}(t) = 1 - e^{-\Lambda t},$$

so that selection is given by:

$$\mu^{Calvo}(t) = \frac{e^{-\Lambda t}}{e^{-\Lambda t}} - 1 = 0.$$

Thus, under Calvo pricing there is no selection. In other words, price changing firms are a representative draw from the population.

4.2 Hazard functions

The empirical literature on price-setting has devoted substantial effort to estimating the shape of the hazard function of price adjustment.[7] The motivation is that, at least since the work of Dotsey, King, and Wolman (1997, 1999), and Wolman (1999), it has been clear that the shape of the hazard function matters for the real effects of monetary shocks.

[6] Vavra (2010) provides a different proof of the fact that the real effects under Taylor are minimum. Without linking it to selection, we first proved that result in Carvalho and Schwartzman (2008).

[7] For a recent review of this literature, see Klenow and Malin (2010, section 5.3).

Assuming G is differentiable, the hazard function can be defined as:

$$h(s) = \frac{\frac{\partial G(s)}{\partial t}}{1 - G(s)}.$$

We start by showing that the concept of selection and hazard functions are closely related. Specifically, the following holds:

Lemma 4. *Let μ and h be, respectively, the selection function and hazard function associated with a differentiable c.d.f. G. Let $t_1 \in [0, \infty)$ be the smallest value of t such that $\omega(t_1) = 1$. Then, for $t < t_1$,*

$$\mu(t) = \int_t^{t_1} \frac{h(s)}{\Lambda} \Psi_t(s)\, ds - 1, \tag{15}$$

where

$$\Psi_t(s) \equiv \frac{1 - G(s)}{\int_t^{t_1} (1 - G(v))\, dv}.$$

is the density of prices of age s among all prices older than t.

Thus, up to a constant, selection at t is proportional to a weighted average of the hazard function evaluated at t and later. The intuition is as follows: Selection at t is tied to the probability of prices set *before* time 0 changing at t. Given stationarity, this is equivalent to the probability of prices changing at age t or *afterwards.* Since the hazard function is the continuous-time analogue of the probability of a price changing at a given age, conditional on it having survived up to that age, selection at date t can be obtained from integrating the hazard function from t to infinity using $\Psi_t(s)$ as weights. The normalization by the average frequency of price changes Λ reflects the fact that, unlike the hazard function, selection does not depend on the average frequency of price changes.

From (15), we can show that if a hazard function is strictly increasing, then there is positive selection at all t:

Lemma 5. *For a given distribution of price durations $G(t)$, consider the corresponding hazard $h(t) = \frac{\partial G(t)}{\partial t} / (1 - G(t))$ and selection $\mu(t) = \frac{1-G(t)}{1-\omega(t)} - 1$ functions. If $h(t') > h(t)$ for all $t' > t$, then $\mu(t) > 0$ for all $t > 0$.*

The result is intuitive. An increasing hazard function implies positive selection, since the probability of a price change increases with the age of the price. An immediate implication is that any economy featuring an increasing hazard of price adjustment will feature higher selection than an economy featuring Calvo pricing. The Lemma thus verifies the intuition spelled out by Wolman (1999) for the reason why, as compared to Calvo pricing, increasing hazard functions are associated with smaller real effects of monetary shocks.

The general intuition behind Lemma 5 extends to the comparison of two hazard functions. In this case, the c.d.f.'s can be ranked in terms of the associated cumulative selection. Given two economies, one with a more increasing hazard function than the other, the economy with the more increasing hazard function features higher cumulative selection and lower monetary non-neutrality:[8]

Proposition 3. *For two economies A and B with the same average frequency of price changes* ($\Lambda_A = \Lambda_B$) *and for which the relevant moments and derivatives are defined, if either*

1) there is a single crossing at some t^* *so that* $h_A(t) \geq h_B(t)$ *for* $t \leq t^*$ *and* $h_A(t) < h_B(t)$ *for* $t > t^*$*, or*

2) $\frac{\partial h_A(t)}{\partial t} < \frac{\partial h_B(t)}{\partial t}$ $\forall t$,

Then $\Xi_A(t) < \Xi_B(t)$ $\forall t$.

Given Proposition 3, it follows immediately from Proposition 2 that monetary shocks are associated with smaller real effects in economies in which the hazard function increases more quickly.

4.3 Heterogeneity in price stickiness

In this section we show that selection effects also shed light on, and allow us to generalize the findings in Carvalho (2006), that a one-sector model calibrated to the average frequency of price changes is likely to understate the real effects of nominal shocks relative to a model with cross-sectoral heterogeneity in price stickiness. These findings are of particular importance because, as documented by Bils and Klenow (2004) and others, there is substantial heterogeneity in the frequency of price changes.

Consider a heterogeneous sticky-price economy with K sectors indexed by k, each with a measure Φ_k of firms and sector-specific distribution of price-durations $G_k(t)$.[9] For notational convenience, we use $E[\cdot]$ to denote cross-sectoral weighted averages:

$$E[x_k] \equiv \sum_{k=1}^{K} \Phi_k x_k. \tag{16}$$

The price level in the heterogeneous economy is:

$$p(t) = E[p_k(t)],$$

[8] The result is actually stronger than this, as all that is required is a single-crossing condition on the two hazards (see the proof in the Appendix).

[9] For brevity we do not specify the whole multisector model here, and borrow the required log-linear equations directly from Carvalho and Schwartzman (2008).

where $p_k(t)$ is the price level in sector k. These sectoral price levels are aggregates of past pricing decisions:

$$p_k(t) = \int_{-\infty}^{t} \Lambda_k \left[1 - G_k(t-s)\right] x_k(s)\, ds,$$

where $\Lambda_k \equiv \left[\int_0^\infty (1 - G_k(s))\, ds\right]^{-1}$ is the average frequency of price changes in sector k.

Definition 1 for selection does not apply to the heterogeneous economy, but it is possible to construct a natural extension. First, if the fraction of new prices in sector k at time t is $\omega_k(t)$, it follows that the fraction of new prices in the economy as a whole (which we denote by $\omega^{het}(t)$) is just the average of the fraction of new prices across sectors:

$$\omega^{het}(t) = E\left[\omega_k(t)\right].$$

Calculating the fraction of new prices among changing prices is slightly more involved. Here, we have to take into account that the mass of prices changing in a given sector at any given interval dt is given by $\Phi_k \Lambda_k dt$ – the mass of firms in the sector, Φ_k, multiplied by the frequency of price changes Λ_k and by the length of time dt. If we denote the economy-wide fraction of new prices among changing prices at time t by $G^{het}(t)$, then:

$$G^{het}(t) = E\left[\frac{\Lambda_k}{E\left[\Lambda_k\right]} G_k(t)\right].$$

We can now generalize Definition 1 to the heterogeneous economy:

Definition 3. *For all t such that $E\left[\omega_k(t)\right] < 1$, selection (at t), denoted by $\mu^{het}(t)$, is defined as*

$$\mu^{het}(t) \equiv \frac{1 - G^{het}(t)}{1 - \omega^{het}(t)} - 1. \tag{17}$$

For t such that $\omega^{het}(t) = 1$,

$$\mu^{het}(t) = 0.$$

Definition 2 also generalizes to the heterogeneous economy in the natural way, so that $\Xi^{het}(t) = \int_0^t \mu^{het}(s)\, ds$. Given those definitions, it is possible to extend Proposition 2 to heterogeneous economies:

Proposition 2'. *Consider the real effects of a shock to nominal aggregate demand given by $m^{new}(t) \geq m^{old}$ for all t in two economies, A and B, characterized by sector specific distribution of price durations $\left\{G_k^A(t)\right\}_{k=1}^{KA}$ and $\left\{G_k^B(t)\right\}_{k=1}^{KB}$ and by sectoral weights $\left\{\Phi_k^A\right\}_{k=1}^{KA}$ and $\left\{\Phi_k^B\right\}_{k=1}^{K}$. Suppose, moreover, that the cross-sectoral average of the frequencies of price changes in both economies is the same, that is, $E\left[\Lambda_k^A\right] = E\left[\Lambda_k^B\right]$. Then, $\Gamma_A < \Gamma_B$ if either*

1) $\mu_A^{het}(t) \geq \mu_B^{het}(t)\ \forall t$, *or*

2) $\Xi_A^{het}(t) \geq \Xi_B^{het}(t)\ \forall t$.

We are now ready to show the role of selection in generating the results in Carvalho (2006). In Carvalho all sectors feature Calvo pricing, with different hazards of price adjustment. This is a particular example of economies where the relevant source of heterogeneity across sectors in the distribution of price durations is summarized by a single sector-specific scaling parameter.

Specifically, let the c.d.f. of price durations in sector k be given by $G_k(t) = \bar{G}(\Lambda_k t)$, with $\int_0^\infty \left(1 - \bar{G}(t)\right) dt = 1$. Note that $\bar{G}$ is a generic c.d.f. common to all sectors, but that the average frequency of price change in sector k is equal to Λ_k.

Given this parametrization of the heterogeneous economy, we can compare it to a counterfactual one-sector economy with c.d.f. of price durations $\bar{G}(E[\Lambda_k]\, t)$, defined below:

Definition 4. *Consider a multisector economy characterized by sector specific distribution of price durations* $\left\{\bar{G}(\Lambda_k t)\right\}_{k=1}^K$, *where* $\int_0^\infty \left(1 - \bar{G}(t)\right) dt = 1$, *and sectoral weights* Φ_k. *The counterfactual one-sector economy is an economy with one sector and c.d.f. of price durations given by* $\bar{G}(E[\Lambda_k]\, t)$.

The following proposition compares the cumulative selection function in both economies:

Proposition 4. *Let* $\Xi^{het}(t)$ *denote cumulative selection of a multisector economy characterized by the sectoral c.d.f.'s of price durations* $\left\{\bar{G}(\Lambda_k t)\right\}_{k=1}^K$, *where* $\int_0^\infty \left(1 - \bar{G}(t)\right) dt = 1$, *and sectoral weights* Φ_k, *and let* $\Xi^{count}(t)$ *denote cumulative selection of its counterfactual one-sector economy. Then,*

$$\Xi^{het}(t) < \Xi^{count}(t)\ \ \forall t. \tag{18}$$

Thus, cumulative selection in the multisector economy is always smaller than in the counterfactual one-sector economy. It follows immediately from Proposition 2' that a shock to nominal aggregate demand in the multisector economy has larger real effects than in the counterfactual one-sector economy.

The intuition for Proposition 4 is easiest to understand in the case considered by Carvalho (2006), where the hazard of price adjustment is constant within each sector, as in Calvo (1983). In this economy, $\bar{G}(t) = 1 - e^{-t}$, so that:

$$G_k(t) = 1 - e^{-\Lambda_k t},\ \omega_k(t) = 1 - e^{-\Lambda_k t}. \tag{19}$$

Each sector features Calvo pricing so that within each sector there is no selection. This, however, is not true in the aggregate. We can check that selection in the heterogeneous economy is negative:

$$\mu^{het}(t) = \frac{\frac{E\left[\Lambda_k e^{-\Lambda_k t}\right]}{E[\Lambda_k]}}{E\left[e^{-\Lambda_k t}\right]} - 1 = \frac{cov\left(\Lambda_k, 1-\omega_k(t)\right)}{E\left[\Lambda_k\right] E\left[1-\omega_k(t)\right]} < 0, \tag{20}$$

where *cov* denotes the cross-sectional covariance given sectoral weights Φ_k.

The covariance term in equation (20) neatly summarizes the intuition behind the main result in this section. In the heterogeneous economy, price changes are disproportionately selected from sectors with high frequency of price changes (high Λ_k). However, exactly because these sectors have a high frequency of price changes, they have a smaller fraction of old prices (low $1-\omega_k(t)$). Therefore, $cov\left(\Lambda_k, 1-\omega_k(t)\right) < 0$ and selection is negative. In contrast, the counterfactual one-sector economy is just a Calvo economy, so that selection is zero. Thus, the heterogeneous economy features lower selection than its counterfactual one-sector counterpart, and higher real effects of monetary shocks.

4.4 The variance of price durations

For economists trying to calibrate time-dependent sticky-price models, the results presented so far may seem a bit discouraging. They imply that the average frequency of price changes is far from a being a sufficient statistic for the real effects of nominal shocks. Rather, they suggest that one cannot do without the whole distribution of price durations, since it is the shape of that distribution that determines selection.[10]

Our next results show that it may not be necessary to account for the entire distribution of price durations. They make the case that, for a given frequency of price changes, the variance of price durations may be a good scalar metric of selection effects and, in some particular cases, a sufficient statistic.

As a first step, we compare selection among two distributions of price durations where one is a mean preserving spread of the other. Proposition 5 states the result:

Proposition 5. *Consider two economies, A and B, characterized by the distribution of price spells G_A and G_B, where G_A is obtained from a mean preserving spread of G_B. Then, $\Xi_A(t) \leq \Xi_B(t)$ $\forall t$. Conversely, if $\Xi_A(t) \leq \Xi_B(t)$ and $\Lambda_A = \Lambda_B$, then G_A is a mean preserving spread of G_B.*

Thus, if we restrict ourselves to comparing economies that can be ordered in terms of cumulative selection, the variance of price durations is a sufficient statistic for that ordering. Given that

[10] For a model calibrated with microeconomic estimates of the full distribution of the duration of price spells, see Vavra (2010). As shown in that paper, an alternative to our approach is to consider the distribution of remaining durations of prices in place. Given stationarity, one is just a transformation of the other.

restriction, in the context of Proposition 5, if the variance of price durations in economy A is higher than in economy B, then, selection is lower in A than in B.

For a given frequency of price changes, the variance of price durations is, furthermore, a sufficient statistic for the real effects of nominal disturbances in the case of shocks to the level of nominal income discussed in Section 3.2.1:[11]

Proposition 6. *Suppose an economy is characterized by a distribution of price spells G with finite mean and variance given by* Λ^{-1} *and* σ^2*. The real effects of a permanent level shock to nominal aggregate demand of size* Δm *satisfy*

$$\lim_{\rho \to 0} \frac{\Gamma}{\Delta m} = \frac{1}{2}\left(\Lambda^{-1} + \Lambda \sigma^2\right). \tag{21}$$

Note that, unlike Proposition 5, Proposition 6 does not require the economies under comparison to be ordered by degree of selection. In that sense, it applies more broadly than previous results that hinged on an ordering by cumulative selection.

Furthermore, the Proposition presents a closed-form expression for the real effects of the monetary shock. As an example, it allows us to easily calculate the real effects of a level shock under Taylor and Calvo pricing. They are given by $\frac{\Lambda^{-1}}{2}\Delta m$ for Taylor and $\Lambda^{-1}\Delta m$ for Calvo. Hence, the real effects of a level shock are twice as large under Calvo pricing than under Taylor pricing.

The variance is not a sufficient statistic for more complicated shocks. Proposition 6 is a special case of Proposition 6', which applies to any shock whose impulse response function can be well approximated by a polynomial function.

Proposition 6'. *Suppose an economy is characterized by a distribution of price spells G with finite moments of order between* 1 *and* $K+1$*. Let the random variable* τ *be the realized duration of price spells. The real effects of a monetary shock characterized by* $m^{new}(t) - m^{old} = \sum_{k=1}^{K} a_k t^{k-1}$ *are:*

$$\lim_{\rho \to 0} \Gamma = \sum_{k=1}^{K} \frac{a_k}{k(k+1)} \frac{E\left[\tau^{k+1}\right]}{E\left[\tau\right]}.$$

The proposition states that the number of (uncentered) moments necessary to characterize the real effects of a monetary shock increases with the number of polynomial terms necessary to approximate the new trajectory of nominal income. For example, consider the effects of permanent shocks to the growth rate of nominal aggregate demand. Such shocks are relevant for periods of

[11] This is not the case more generally because orderings by variance do not imply an ordering by mean preserving spreads.

disinflation such as the early 80s in the United States. The growth rate shock is:[12]

$$m(t) = \begin{cases} m^{old}, \ t < 0, \\ m^{old} + bt, \ t \geq 0. \end{cases} \tag{22}$$

In that case, the real effects of the shock are given by:

$$\lim_{\rho \to 0} \Gamma = \frac{b}{6} \left(\Lambda^{-2} + 3\sigma^2 + \Lambda \eta \sigma^3 \right),$$

where σ^2 is the variance of price durations and η is the skewness. It is, again, straightforward to calculate the real effects of this kind of shock under Taylor and Calvo pricing. They are, respectively, $\frac{\Lambda^{-1}}{6}b$ and $\Lambda^{-1}b$, so that the real effects of the shock are 6 times as large under Calvo than under Taylor.

5 Interactions in pricing decisions

The analytical results presented above hold under strategic neutrality in price setting. In this section we perform some numerical exercises to assess whether our main results extend to more general cases.

We consider the real effects of a shock across different sticky-price economies indexed by T and characterized by the following family of survival functions:

$$1 - G^T(t) = \begin{cases} 1 - e^{-\theta t} \text{ if } t < T, \\ 0 \text{ if } t \geq T, \end{cases} \tag{23}$$

with θ such that:

$$\int_0^\infty \left(1 - G^T(t)\right) dt = D \text{ for all } T. \tag{24}$$

That is, for different values of T, we adjust θ to ensure that the average duration of price spells equals D. We take the unit of time to be a quarter and set $D = 2$, so that the average price-spell lasts 2 quarters in all economies.

This family includes the two leading cases of constant duration (Taylor, 1979) and constant hazard (Calvo, 1983). The first obtains if $T = 2$ and $\theta = 0$. Calvo pricing obtains with $T \to \infty$ and $\theta \to \frac{1}{2}$.

For any $T'' > T'$, it is easy to check that the survival function parameterized by T'' is a mean preserving spread of the survival function parameterized by T'.[13] Thus, from Proposition 5 it

[12] While this shock involves permanently non-zero inflation, Carvalho (2008, Appendix A.6) shows that, as long as the discount rate (ρ) is not strictly equal to zero, our result is a good approximation for temporary but highly persistent shocks to the growth rate of nominal income, so that inflation converges slowly back to zero.

[13] In particular, it is straightforward to verify that for any T'' and T' with $T'' > T'$, $\int_0^t G^{T''}(s)\,ds \geq \int_0^t G^{T'}(s)\,ds$

follows that, as T increases, cumulative selection decreases.

To perform the simulations, we consider the discrete-time analogue of the model in Section 2. The discrete-time analogue of the family of survival functions described in equations (23) and (24) is:

$$1 - G_t^T = \begin{cases} \theta^t \text{ if } t < T, \\ 0 \text{ if } t \geq T. \end{cases}$$

with θ such that:

$$\sum_{t=0}^{\infty} (1 - G_t)\, dt = 2 \text{ for all } T.$$

The reset price chosen by all firms adjusting in period t is:

$$x_t = \sum (1 - G_t) \left[\alpha m_t + (1 - \alpha) p_t \right], \tag{25}$$

where α determines whether pricing decisions are strategic complements or strategic substitutes. In our experiments we compare results with $\alpha = 1$ (strategic neutrality), $\alpha = 1/3$ (strategic complementarity) and $\alpha = 3$ (strategic substitutability).[14]

In order to perform the numerical exercises, we also need to parameterize the shock process. We follow Mankiw and Reis (2002) and consider a process that is mean reverting in the growth rate of nominal aggregate demand:

$$\Delta m_t = 0.5 \Delta m_{t-1} + \epsilon_t. \tag{26}$$

For given T, equations (25), (26) and the discrete-time analogue of the aggregate price-level equation (6) define a standard linear rational-expectations model in $\{p_t, m_t, x_t\}$, which we solve using Dynare.

Figure 1 shows the log cumulative real effects of monetary shocks described in equation (26) for different levels of strategic interactions, and different T's. As expected, the real effects are, for a given T, largest under strategic complementarity ($\alpha = \frac{1}{3}$) and smallest under strategic substitutability ($\alpha = 3$). Furthermore, for given α, they also increase noticeably as selection decreases (T increases).

Lastly, note that moving from a model with Taylor pricing ($T = 2$ and $\theta = 0$) to one approaching a constant hazard of price adjustment ($T = 20$ and $\theta = 0.4998$) implies an increase in the real effects by a factor of approximately two under strategic neutrality ($\alpha = 1$). This is close to what is implied by the analytical result in Proposition 6, even though the nominal income process in equation (26) does not imply instantaneous level shifts as assumed in the proposition. With

for all t.

[14] Recall that $\alpha = \frac{\sigma + \psi^{-1}}{1 + \theta \psi^{-1}}$. The parametrization with strategic complementarities obtains if $\sigma = 1$, $\psi = 1$, and $\theta = 5$. The parametrization with strategic substitutability obtains if $\sigma = 3$, $\psi \to \infty$, and any θ.

strategic complementarities, real effects increase by a factor greater than two, whereas with strategic substitutability, they increase by a factor smaller than two.

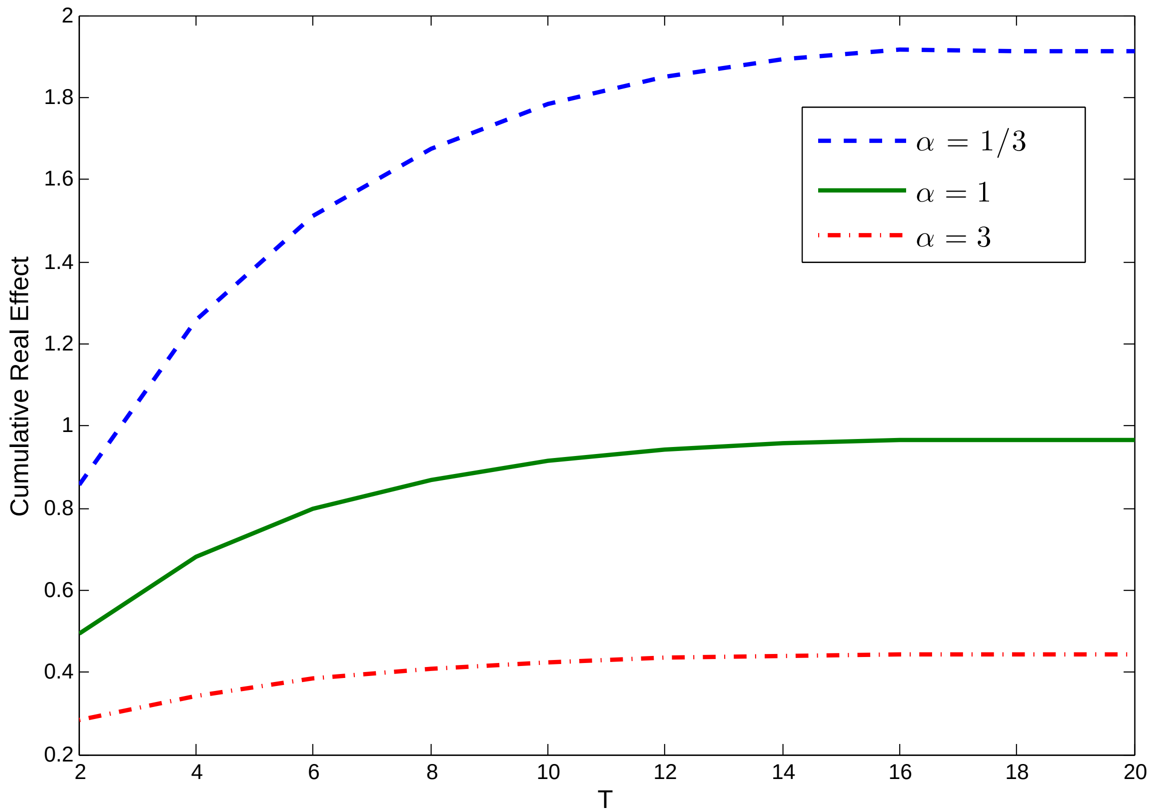

Figure 1: Cumulative real effect a shock to the level of nominal income: Same average duration, different variances.
Note: Average price duration is equal to 2. Distribution of price durations follow equations (23) and (24), see Appendix for details.

6 Summary and conclusion

We investigate the different ways in which the shape of the distribution of duration of price spells affects the real effects of nominal aggregate demand shocks. We highlight a mechanism that so far has barely been given attention in the literature: a selection for the time in which prices were last adjusted. In fact, we show that selection provides a complete characterization of the distribution of price durations in time-dependent sticky-price models.

The results in the paper suggest that a careful characterization of the distribution of price durations is of crucial importance for the proper evaluation of the real aggregate effect of monetary shocks. While the results are derived for the case of time-dependent pricing, there is no reason why the selection effect identified here should not hold some relevance more broadly, whenever the timing of price changes is not entirely up to the discretion of the firms. This suggests that further research on price setting would do well to focus on models that are able to fully account for the distribution price spells and investigating the extent to which the mechanisms emphasized here continue to matter.

References

[1] Alvarez, F., F. Lippi, and L. Paciello (2012), “Monetary Shocks in a Model with Inattentive Producers,” mimeo.

[2] Alvarez, F., H. Le Bihan and F. Lippi (2014), “Small and Large Price Changes and the Propagation of Monetary Shocks”, NBER Working Paper 20155.

[3] Bomono, M. and C. Carvalho (2004) “Endogenous time-dependent rules and inflation inertia.” *Journal of Money, Credit and Banking*: Vol. 36, N. 6, pp. 1015-1041.

[4] Benigno, P. and M. Woodford (2005), “Inflation Stabilization and Welfare: The Case of a Distorted Steady State,” *Journal of the European Economic Association* 3: 1185-1236.

[5] Bils, M. and P. Klenow (2004), “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy* 112: 947-985.

[6] Caballero, R. (1989), “Time dependent rules, aggregate stickiness and information externalities,” *Columbia University Working Paper.*

[7] Calvo, G. (1983), “Staggered Prices in a Utility Maximizing Framework,” *Journal of Monetary Economics* 12: 383-98.

[8] Caplin A. and D. Spulber (1987), “Menu Costs and the Neutrality of Monery,” *Quarterly Journal of Economics* 102: 703-726.

[9] Carvalho, C. (2006), “Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks,” *Frontiers of Macroeconomics*: Vol. 2 : Iss. 1, Article 1.

[10] ____________ (2008), *Heterogeneity in Price Setting*, Ph.D. dissertation, Princeton University.

[11] Carvalho, C. and F. Schwartzman (2008), “Heterogeneous Price Setting Behavior and Aggregate Dynamics: Some General Results,” mimeo available at http://cvianac.googlepages.com/papers.

[12] Chari, V.V., P. Kehoe and E. McGrattan (2000), “Sticky price models of the business cycle: can the contract multiplier solve the persistence problem?” Econometrica 68.5: 1151-1179.

[13] Dotsey, M., R. King and A. Wolman (1997), “State-Dependent Pricing and Dynamics of Business Cycle” Federal Reserve Bank of Richmond Working Paper Series, No. 97-2.

[14] ____________ (1999), "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics* 114: 655-690.

[15] Kiley, M. (2002), "Partial Adjustment and Staggered Price Setting," *Journal of Money, Banking and Credit* 34: 283-298.

[16] Klenow, P., and O. Kryvtsov (2008), "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *Quarterly Journal of Economics* 123: 863-904.

[17] Klenow, P. and B. Malin (2010), "Microeconomic Evidence on Price Setting," *Handbook of Monetary Economics,* Ed. 1, vol. 16(6): 231 - 284.

[18] Mankiw, G. and R. Reis (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics* 117: 1295-1328.

[19] Reis, R. (2006), "Inattentive Producers", *The Review of Economic Studies* 73 (3): 793-821.

[20] Sheedy (2010), "Intrinsic Inflation Persistance," *Journal of Monetary Economics* 57: 1049-1061.

[21] Taylor, J. (1979), "Staggered Wage Setting in a Macro Model," *American Economic Review* 69: 108-113.

[22] Vavra, J. (2010), "The Empirical Price Duration Distribution and Monetary Non-Neutrality," mimeo available at https://sites.google.com/site/jvavra/docs/paper10-29-10_fordistribution.pdf?attredirects=0.

[23] Wolman, A. (1999), "Sticky Prices, Marginal Cost, and the Behavior of Inflation," Federal Reserve Bank of Richmond Economic Quarterly 85 (Fall 1999), 29-48.

[24] Wooodford, M. (2001), "Inflation Stabilization and Welfare," *Contributions to Macroeconomics*: Vol. 2: Iss.1, Article 1.

[25] ____________(2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

[26] ____________(2014), "A Note on Increasing Hazard Functions and the Monetary Transmission Mechanism", *Macroeconomic Dynamics,* Forthcoming.

A Appendix

A.1 A welfare-based measure of real effects

We measure the degree of monetary non-neutrality by the discounted cumulative effect of the shock on the output gap:

$$\Gamma = \int_0^\infty e^{-\rho t} \left[y_1 \left(t \right) - y_0 \right] dt,$$

where ρ is the discount rate and y_0 is the counter-factual path for output that would have held if the shock had never happened. Here we show that, up to a first-order approximation, Γ is proportional to the total impact of the shock on the representative agent's utility. To see this, recall that the utility function is:

$$U \left(t \right) = \int_t^\infty e^{-\rho s} \left[\frac{C \left(t+s \right)^{1-\sigma}}{1-\sigma} - \frac{\int_0^1 L_j \left(t+s \right)^{1+\frac{1}{\psi}} dj}{1+\frac{1}{\psi}} \right],$$

where

$$C \left(t+s \right) \equiv \left[\int_0^1 C_j \left(t+s \right)^{\frac{\varepsilon-1}{\varepsilon}} dj \right]^{\frac{\varepsilon}{\varepsilon-1}}.$$

Taking the total derivative of $U(t)$ with respect to $\ln \left(C_j \left(t+s \right) \right)$ and $\ln \left(L_j \left(t+s \right) \right)$ yields the first-order approximation:

$$u \left(t \right) = \int_t^\infty e^{-\rho s} \left(C^{1-\sigma} \int_0^1 c_j \left(t+s \right) dj - L^{1+\frac{1}{\psi}} \int_0^1 l_j \left(t+s \right) dj \right) dt,$$

where $u(t)$ is the *linear* approximation of the utility function as measured from $t = 0$ onward, C is steady-state consumption, L is steady-state aggregate hours worked and, as before, $c_j \left(t+s \right)$ and $l_j \left(t+s \right)$ are *log-linear* deviations from steady state of consumption of different varieties and employment at different firms, respectively.

In equilibrium, $c_j \left(t+s \right) = l_j \left(t+s \right) = y_j \left(t+s \right)$. Also, up to a first-order approximation, $y \left(t \right) = \int_0^1 y_j \left(t+s \right) dj$. Finally, in steady state, $C = L = Y$. Hence,

$$u \left(t \right) = \left(Y^{1-\sigma} - Y^{1+\frac{1}{\psi}} \right) \int_t^\infty e^{-\rho s} y \left(t+s \right) dt.$$

When comparing the utility under two different trajectories for nominal aggregate demand we can write:

$$\begin{aligned} u_1 \left(0 \right) - u_0 \left(0 \right) &= \left(Y^{1-\sigma} - Y^{1+\frac{1}{\psi}} \right) \int_0^\infty e^{-\rho t} \left[y_1 \left(t \right) - y_0 \right], \\ &= \left(Y^{1-\sigma} - Y^{1+\frac{1}{\psi}} \right) \Gamma. \end{aligned}$$

Thus, the utility impact of the shock is, up to a first-order approximation, proportional to Γ. It remains to prove that the first term, in brackets, is greater than zero, to be sure the effect goes in the right direction.

Optimal price-setting in the zero inflation steady state implies that:

$$P = \frac{\varepsilon}{\varepsilon - 1} W = \frac{\varepsilon}{\varepsilon - 1} P L^{\frac{1}{\psi}} C^{\sigma} = \frac{\varepsilon}{\varepsilon - 1} P Y^{\frac{1}{\psi}+\sigma},$$

where the second equation follows from the optimality condition for labor supply (equation 1) and the third equality follows from the production technology, the goods market clearing condition and the assumption of a symmetric steady-state equilibrium. It follows that:

$$Y = \left(1 - \frac{1}{\varepsilon}\right)^{\frac{1}{\sigma + \frac{1}{\psi}}},$$

and, since $\varepsilon > 1$, $\sigma > 0$ and $\psi > 0$,

$$Y^{1-\sigma} - Y^{1+\frac{1}{\psi}} > 0.$$

At this point it is important to pause for a comparison with the literature, in particular Woodford (2001) and Benigno and Woodford (2005). First, in contrast to Woodford (2001), we do not include in the model a subsidy that undoes the monopolistic distortion. This generates a first-order positive utility impact of a surprise increase in nominal aggregate demand. The reason is that, by surprising firms, the nominal shock reduces their markup leading to more efficient production.

The second point is that, in contrast with both Woodford (2001) and Benigno and Woodford (2005), we only approximate the utility function up to first order. By doing this we miss the component of the welfare cost of inflation most emphasized in these papers: the ensuing dispersion in prices that distorts allocation across varieties of products. Thus, our results do not speak to the impact of nominal aggregate demand shocks on this component of welfare.

Lastly, we do not mean to imply that we have a superior welfare criterion to analyze optimal policy than Woodford (2001) or Benigno and Woodford (2005), or that our results suggest that surprise positive nominal shocks are a commendable policy. In particular, Benigno and Woodford (2005) show that even in the presence of monopolistic distortions, for most reasonable parameterizations, optimal policy under commitment should concern itself primarily with price stabilization and should not attempt to undo the monopolistic distortion by surprising firms. Still, we contend that to the extent that monetary surprises do take place, it is important to understand how they impact household's welfare, and Γ provides a useful measure of this impact.

A.2 The sticky-information model

We now describe a sticky-information model analogous to the sticky-price model in the main text. The comparison with this model plays a key role in the proof of Proposition 1 which, as we show, can be alternatively formulated as stating that the real effects of a monetary shock in a sticky-price model are identical to those effects in a sticky-information model, so long as the distribution of price spells in the former is identical to the distribution of price plans in the latter. One immediate implication of the proposition thus formulated is that all the results in the paper apply equally to time-dependent variants of Mankiw and Reis (2002).

We lay out a continuous-time version of Mankiw and Reis (2002) with general survival functions for price plans. The household and the market structure are identical to the one laid out in Section 2. Firms are also identical, except that instead of choosing prices that remain in place for a period of time, firms choose *price plans*, which are insensitive to new information. In particular, under sticky information, a firm that resets its price *plan* $X_j(t,s)$ at time t solves:

$$\max_{X_j(t)} E_t\left[\int_0^\infty e^{-\rho(s-t)}(1-G(s))\left[X_j(t,s)Y_j(t+s)-W_j(t+s)N_j(t+s)\right]ds\right]$$

$$s.t.\ Y_j(t+s)=N_j(t+s),$$

$$Y_j(t+s)=\left(\frac{X_j(t,s)}{P(t+s)}\right)^{-\varepsilon}Y(t+s),$$

where $N_j(t+s)$ is the amount of labor demanded by the firm, and where the demand function already takes into account that goods market clearing implies $C_j(t)=Y_j(t)$.

Firms operate under perfect foresight, except for the fact that they do not anticipate the shock. Given that, the first order condition implies that:

$$X_j(t,s)=\frac{\varepsilon}{\varepsilon-1}E_t\left[W_j(t+s)\right],$$

where $E_t\left[W_j(t+s)\right]=W_j^{old}(t+s)$ for $t<0$ and $E_t\left[W_j(t+s)\right]=W_j^{new}(t+s)$ for $t\geq 0$.

The monetary shock is specified in the same way as in Section 2. As before, we log-linearize the model around a deterministic, zero-inflation symmetric steady state. In this log-linear environment, the optimal reset price plan for firms that change prices at time t is (lowercase variables denote log-deviations from the steady state):

$$x(t,s)=E_t\left[w_j(t+s)\right]. \tag{27}$$

A sequence of substitutions analogous to the ones laid out in Section 2 yields:

$$x(t,s) = E_t\left[\alpha m(t+s) + (1-\alpha)p(t+s)\right],$$

where $\alpha = \frac{\sigma+\psi^{-1}}{1+\varepsilon\psi^{-1}}$. Finally, the aggregate price level is given by:

$$p(t) = \int_{-\infty}^{t} \Lambda\left(1 - G(t-v)\right) x(v,s)\, dv.$$

In analogy to the sticky-price model, we can partition firms into those with "old" price plans and those with "new" price plans, depending on whether the price plan was set before or after the shock. Under sticky information, all firms with the same information set choose the same price path. Hence, at any given point in time, all firms with old price plans set the same price, and all firms with new price plans also have the same price, irrespective of the path of nominal income.

It follows that, for $\alpha = 1$:

$$p(t) = \omega(t)\, m^{new}(t) + (1-\omega(t))\, m^{old}.$$

The real effects of a nominal shock are (using the fact that $p^{old} = m^{old}$):

$$\begin{aligned} \Gamma^{si} &= \int_0^\infty e^{-\rho t}\left(m^{new}(t) - p^{new}(t) - \left(m^{old} - p^{old}\right)\right) dt = \\ &= \int_0^\infty e^{-\rho t}\left(1-\omega(t)\right)\left(m^{new}(t) - m^{old}\right). \end{aligned}$$

Note that Proposition 1 states that, in a sticky-price economy with strategic neutrality in price setting, $\Gamma = \int_0^\infty e^{-\rho t}\left(1-\omega(t)\right)\left(m^{new}(t) - m^{old}\right)$. Hence, if $\alpha = 1$ and the distribution of durations of price plans in a sticky-information economy is the same as the distribution of price spells in a sticky-price economy, Proposition 1 states that the real effects of a monetary shock in both economies are the same. Importantly, since we can define selection in a sticky-information economy just as in the sticky-price economy, it follows that all the analytical results in the paper concerning the relationship between selection and real effects of nominal shocks are equally valid in a sticky-information environment.

A.3 Proofs

Lemma A. 1. *Let $G(t)$, $\omega(t)$, $\mu(t)$ and $\Xi(t)$ be as defined in Section 3. Then*

$$1 - \omega(t) = e^{-\Lambda t - \Lambda \int_0^t \mu(v) dv} = e^{-\Lambda t - \Lambda \Xi(t)}.$$

Proof of Lemma A.1. The second equation follows from the definition of cumulative selection. We prove that the first equation holds.

Let t_1 be the smallest t such that $\omega(t) = 1$. Since $\omega(t)$ is monotonically increasing, $\omega(t) = 1$ for all $t \geq t_1$ and $\omega(t) < 1$ otherwise.

We consider two cases: $t \in [0, t_1)$ and $t \in [t_1, \infty)$. We use a guess and verify procedure.

1) $t \in [0, t_1)$:

First, note that for $t = 0$, $1 - \omega(t) = 1$ and $e^{-\Lambda t - \int_0^t \mu(v)dv} = 1$. Thus, we confirm that the guess works for $t = 0$. We verify the guess also holds for $t \in [0, t_1)$ if we can show that the derivative of both sides of the equation with respect to t are identical over that range.

The derivative of the left-hand-side is:

$$\frac{\partial(1 - \omega(t))}{\partial t} = -\frac{\partial\omega(t)}{\partial t} = -\Lambda(1 - G(t)),$$

and the derivative of the right-hand-side is:

$$\frac{\partial e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv}}{\partial t} = -\Lambda(1 + \mu(t))\, e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv}$$

Given our guess, $e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv} = 1 - \omega(t)$, so that

$$-\Lambda(1 + \mu(t))\, e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv} = -\Lambda(1 + \mu(t))(1 - \omega(t)).$$

By definition, $1 + \mu(t) = \frac{1-G(t)}{1-\omega(t)}$ so that

$$\frac{\partial e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv}}{\partial t} = -\Lambda(1 - G(t)).$$

2) $t \geq t_1$:

We know that $1 - \omega(t) = 0$ for $t \geq t_1$. We need to show that this also the case for $e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv}$. First, note that, given the result in (1) above, we know that, taking the left-limit as $t \to t_1^-$, we find that $\lim_{t\to t_1^-} e^{-\Lambda t - \Lambda\int_0^t \mu(v)dv} = \lim_{t\to t_1^-}(1 - \omega(t)) = 0$.

Since $1 - \omega(t) = 0$ for $t \geq t_1$, it remains to show that it is also the case that $e^{-\Lambda t - \Lambda, \int_0^t \mu(v)dv} = 0$ for all $t \geq t_1$. From the proposed solution, for any t and t^{prime} such that $t > t'$, $1 - \omega(t) = (1 - \omega(t'))\, e^{-\Lambda(t-t') - \Lambda\int_{t'}^t \mu(v)dv}$. Taking the left-limit again and applying the definition of μ, so that $\mu(t) = 0$ for $t > t_1$, it follows that, for $t > t_1$:

$$
\begin{aligned}
1-\omega(t) &= \lim_{t'\to t_1^-}\left(1-\omega\left(t'\right)\right)e^{-\Lambda(t-t')-\Lambda\int_{t'}^{t}\mu(v)dv}\\
&= \lim_{t'\to t_1^-}\left(1-\omega\left(t'\right)\right)e^{-\Lambda(t-t')}=0.
\end{aligned}
$$

Given that $e^{-\Lambda t-\Lambda,\int_0^t \mu(v)dv}=0$ for all $t>t_1$, it follows that $\lim_{t'\to t_1^+} e^{-\Lambda t-\Lambda,\int_0^t \mu(v)dv}=0$. Thus, both the left and right limits coincide and $e^{-\Lambda t_1-\Lambda,\int_0^{t_1}\mu(v)dv}=0$.

□

Proof of Lemma 1. We check that f is given by

$$
G(t)=1-(1+\mu(t))e^{-\Lambda t-\Lambda\int_0^t \mu(s)ds}\ \forall t.
$$

Integrating equation (8) we find that

$$
1-\omega(t)=1-\Lambda\int_0^t(1-G(s))\,ds.
$$

From the sequence of equalities in (9) it follows that

$$
1-\Lambda\int_0^t(1-G(s))\,ds=1-\omega(t)=e^{-\Lambda t-\Lambda\int_0^t\mu(s)ds}.
$$

Thus, we can write

$$
G(t)=1-(1+\mu(t))\left[1-\Lambda\int_0^t(1-G(s))\,ds\right].
$$

Use the definition of $\mu(t)$ to substitute it out, to get, for t such that $\omega(t)<1$

$$
\begin{aligned}
G(t) &= 1-\left(1+\frac{1-G(t)}{1-\Lambda\int_0^t(1-G(s))\,ds}-1\right)\left[1-\Lambda\int_0^t(1-G(s))\,ds\right]\\
&= G(t).
\end{aligned}
$$

For t such that $\omega(t)=1$,

$$
G(t)=1-(1-1)\left[1-\Lambda\int_0^t(1-G(s))\,ds\right]=1 \tag{28}
$$

Since $\omega(t)=1-\Lambda\int_0^t(1-G(s))\,ds$, it follows that if $\omega(t)=1$, $\int_0^t(1-G(s))\,ds=\Lambda^{-1}$. But, by definition, $\Lambda^{-1}=\int_0^\infty(1-G(s))\,ds$. Hence, if $\omega(t)=1$, $G(t)=1$, as implied by equation (28).

□

Proof of Lemma 2. 1) and 2) follow from inspection of equation (12). □

Proof of Proposition 1. Let $\Gamma^{sp}\left(m^{old}, m^{new}(t); G\right)$ and $\Gamma^{si}\left(m^{old}, m^{new}(t); G\right)$ be the cumulative real effects of an unexpected change in the path of nominal aggregate demand from m to $m^{new}(t)$ in, respectively, a sticky price as defined in the text and in a sticky information economy as defined in Appendix A.2, with G the c.d.f that summarizes the arrival of adjustment opportunities in each of these economies. We show that, if $\alpha = 1$, $\Gamma^{sp}(m_0(t), m_1(t); G) = \Gamma^{si}(m_0(t), m_1(t); G) = \int_0^\infty (1-\omega(t))\left(m^{new}(t) - m^{old}\right) dt$, where the second equality follows from the discussion in Appendix A.2. We denote the "new" and "old" paths of the various endogenous variables in the sticky information and sticky price economies by $y^{si,new}(t)$, $y^{si,old}$, $y^{sp,new}(t)$, $y^{sp,old}$, etc.

Define $\Delta_\Gamma \equiv \Gamma^{si} - \Gamma^{sp}$. Then:

$$\begin{aligned}
\Delta_\Gamma &= \int_0^\infty e^{-\rho t}\left(y^{si,new}(t) - y^{si,old}\right) dt - \int_0^\infty e^{-\rho t}\left(y^{sp,new}(t) - y^{sp,old}\right) dt = \\
&= \int_0^\infty e^{-\rho t}\left(m^{new}(t) - p^{si,new}(t) - \left(m^{old} - p_0^{si,old}\right)\right) dt \\
&\quad - \int_0^\infty e^{-\rho t}\left(m^{new}(t) - p^{sp,new}(t) - \left(m^{old} - p^{sp,old}\right)\right) dt = \\
&= \int_0^\infty e^{-\rho t}\left(\left[p^{sp,new}(t) - p^{sp,old}\right] - \left[p^{si,new}(t) - p^{si,old}\right]\right) dt.
\end{aligned}$$

Since m^{old} is a constant, it follows that $p^{si,old} = p^{sp,old} = m^{old}$ and

$$\Delta_\Gamma = \int_0^\infty e^{-\rho t}\left(p^{sp,new}(t) - p^{si,new}(t)\right).$$

Recall that

$$\begin{aligned}
p^{sp,new}(t) &= \int_0^t \Lambda(1 - G(t-v))\, x^{sp,new}(v)\, dv + \int_{-\infty}^0 \Lambda(1 - G(t-v))\, x^{sp,old}(v)\, dv \\
&= \int_0^t \Lambda(1 - G(t-v))\, x^{sp,new}(v)\, dv + p_0^{sp,old} = \\
&= \int_0^t \Lambda(1 - G(t-v))\, x^{sp,new}(v)\, dv + m_0,
\end{aligned}$$

and, analogously,

$$p^{si,new}(t) = \int_0^t \Lambda(1 - G(t-v))\, x^{si,new}(v,t)\, dv + m_0,$$

so that:

$$\Delta_\Gamma = \int_0^\infty e^{-\rho t}\left(\left[\int_0^t \Lambda\left(1-G\left(t-v\right)\right)\left[x^{sp,new}\left(v\right)-x^{si,new}\left(v,t\right)\right]dv\right]\right)dt.$$

We prove that $\int_0^\infty e^{-\rho t}\int_0^t \Lambda\left(1-G\left(t-v\right)\right)\left[x^{sp,new}\left(v\right)-x_1^{si,new}\left(v,t\right)\right]dvdt=0$. First write the integral as:

$$\int_0^\infty\int_0^\infty \mathbb{1}\left(v\leq t\right)e^{-\rho t}\Lambda\left(1-G\left(t-v\right)\right)\left[x^{sp,new}\left(v\right)-x^{si,new}\left(v,t\right)\right]dvdt,$$

where $\mathbb{1}\left(.\right)$ is the indicator function. Now do the substitution: $v=z$ and $t=z+w$:

$$\int_0^\infty\int_0^\infty e^{-\rho(z+w)}\mathbb{1}\left(z\leq z+w\right)\Lambda\left(1-G\left(w\right)\right)\left[x^{sp,new}\left(z\right)-x^{si,new}\left(z,z+w\right)\right]dwdz$$

Note that the indicator function is always equal to 1 for all z and $w\geq 0$, so we can write:

$$\int_0^\infty\int_0^\infty e^{-\rho(z+w)}\Lambda\left(1-G\left(w\right)\right)\left[x^{sp,new}\left(z\right)-x^{si,new}\left(z,z+w\right)\right]dwdz$$

Now, we show that the inner integral is equal to zero. We can take all multiplicative terms that only depend on z out of the inner integral and rearrange it to get

$$\begin{aligned}
&\int_0^\infty e^{-\rho(z+w)}\Lambda\left(1-G\left(w\right)\right)\left[x^{sp,new}\left(z\right)-x^{si,new}\left(z,z+w\right)\right]dw=\\
=\;&\Lambda e^{-\rho z}\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)dw\times x^{sp,new}\left(z\right)-\Lambda e^{-\rho z}\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)x^{si,new}\left(z,z+w\right)dw=\\
=\;&\Lambda e^{-\rho z}\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)dw\times\left[x^{sp,new}\left(z\right)-\frac{\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)x^{si,new}\left(z,z+w\right)dw}{\int_0^\infty e^{-\rho(z+w)}\Lambda\left(1-G\left(w\right)\right)dw}\right].
\end{aligned}$$

Optimal price-setting implies:

$$\begin{aligned}
x^{sp,new}\left(z\right) &= \frac{\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)m^{new}\left(z+w\right)dw}{\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)dw},\\
x^{si,new}\left(z,z+w\right) &= m^{new}\left(z+w\right),
\end{aligned}$$

so that

$$x^{sp,new}\left(z\right)=\frac{\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)x^{si,new}\left(z,z+w\right)dw}{\int_0^\infty e^{-\rho w}\left(1-G\left(w\right)\right)dw}.$$

It follows that

$$\int_0^\infty\int_0^t \Lambda\left(1-G\left(t-v\right)\right)\left[x^{sp,new}\left(v\right)-x^{si,new}\left(v,t\right)\right]dv=0.$$

□

Proof of Proposition 2. 1) and 2) follow from inspection of equation (13) and from the fact that 1) implies 2). □

Proof of Lemma 3. Consider any alternative pricing rule G with the average frequency of price changes Λ. If G(t)=1, $\mu(t) = 0$ and the result follows trivially. Otherwise, we denote the corresponding selection at t as $\mu(t)$:

$$\mu(t) = \frac{1 - G(t)}{1 - \omega(t)} - 1. \tag{29}$$

The proof of the result follows from a comparison of numerators and denominators in (14) and (29) for a given Λ. First, it is immediate that $1 - G(t) \leq 1$. Second, we can show that $1 - \omega(t) \geq 1 - \Lambda t$:

$$1 - \omega(t) = 1 - \Lambda \int_0^t (1 - G(s))\, ds = 1 - \Lambda t + \int_0^t G(s)\, ds \geq 1 - \Lambda t.$$

Thus, wherever $\mu^{Taylor}(t)$ is positive, $\mu^{Taylor}(t) \geq \mu(t)$ for any G.

Cumulative selection under Taylor pricing is:

$$\Xi^{Taylor}(t) = \begin{cases} -\frac{\ln(1-\Lambda t)}{\Lambda} - t \text{ if } t < \Lambda^{-1}, \\ \infty \text{ otherwise.} \end{cases}$$

Since under Taylor pricing selection is the highest possible for $t < \Lambda^{-1}$, and for $t \geq \Lambda^{-1}$ cumulative selection is infinite, it follows that cumulative selection is the highest possible everywhere.

□

Proof of Lemma 4. We verify that the statement is true through a sequence of substitutions. Substituting out $\Psi_t(s)$ in the right-hand-side, we get that $\mu(t) = \int_t^{t_1} \frac{h(s)}{\Lambda} \frac{1-G(s)}{\int_t^{\infty}(1-G(v))dv} ds - 1$. From equation (8) it is easy to verify that $\Lambda \int_t^{t_1} (1 - G(v))\, dv = 1 - \omega(t)$. Since it does not depend on s we can take it out of the integral to get $\mu(t) = \frac{1}{1-\omega(t)} \int_t^{t_1} h(s)(1 - G(s))\, ds - 1$. Finally, note that $h(s)(1 - G(s)) = \frac{\frac{\partial G(s)}{\partial s}}{1-G(s)}(1 - G(s)) = \frac{\partial G(s)}{\partial s}$, so that $\mu(t) = \frac{1}{1-\omega(t)} \int_t^{t_1} \frac{\partial G(s)}{\partial s} ds - 1 = \frac{1}{1-\omega(t)} \int_t^{\infty} \frac{\partial G(s)}{\partial s} ds - 1 = \frac{1-G(t)}{1-\omega(t)} - 1$, where the second equality follows from the fact that $G(t) = 1$ for $t > t_1$. This is exactly how $\mu(t)$ is defined in **Definition 1** for $\omega(t) < 1$.

□

Proof of Lemma 5. First, note that $\mu(0) = 0$ always, since $G(0) = \omega(0) = 0$. Second, we can write $\mu(t) = E_{\Psi_0}\left[\frac{h(s)}{\Lambda} | s \geq t\right] - 1$, where the conditional expectation is taken with respect to the probability measure Ψ_0. If the hazard function is (weakly) increasing, it follows that $\mu(t) =$

$E_{\Psi_0}\left[\frac{h(s)}{\Lambda}|h\left(s\right)\geq h\left(t\right)\right]-1>0$. Also, since, by assumption, $h\left(t\right)$ increases in t, so does $\mu\left(t\right)$. Since $\mu\left(0\right)=0$, $\mu\left(t\right)>0\ \forall t>0$. □

Proof of Proposition 3. 1) We prove this in two steps:

i) *There is a unique* $t^{**}>0$ such that $G_A\left(t^{**}\right)=G_B\left(t^{**}\right)<1$, $1-G_A\left(t\right)<1-G_B\left(t\right)$ if $t<t^*$ and $1-G_A\left(t\right)>1-G_B\left(t\right)$ if $t>t^*$.

First we show that a t^{**} with $G_A\left(t^{**}\right)=G_B\left(t^{**}\right)<1$ exists. Suppose not, then $G_A\left(t\right)>G_B\left(t\right)\ \forall t$ or vice versa (otherwise, since G is differentiable, it is continuous and t^{**} must exist by the intermediate point theorem). But this contradicts the assumption that $\Lambda_A=\Lambda_B$, since $\Lambda_A=\int_0^\infty\left(1-G_A\left(t\right)\right)dt$ and $\Lambda_B=\int_0^\infty\left(1-G_B\left(t\right)\right)$.

Second, we show that $t^{**}>t^*$. Note that:

$$\begin{aligned}1-G_A\left(t\right) &= e^{-\int_0^t h_A(s)ds},\\ 1-G_B\left(t\right) &= e^{-\int_0^t h_B(s)ds}.\end{aligned}$$

Since $h_A\left(t\right)>h_B\left(t\right)\ \forall t<t^*$, it follows that

$$1-G_A\left(t\right)<1-G_B\left(t\right)\ \ \forall t<t^*.$$

It follows that $t^{**}>t^*$.

Let t^{**} be the first crossing point. Since $e^{-\int_0^{t^{**}}h_A(s)ds}=e^{-\int_0^{t^{**}}h_B(s)ds}$, we can write

$$\begin{aligned}1-G_A\left(t\right)-\left(1-G_B\left(t\right)\right) &= e^{-\int_0^t h_A(s)ds}-e^{-\int_0^t h_B(s)ds}\\ &= e^{-\int_0^{t^{**}}h_A(s)ds}\left(e^{-\int_{t^{**}}^{t}h_A(s)ds}-e^{-\int_{t^{**}}^{t}h_B(s)ds}\right)\ \text{if}\ t>t^{**}.\end{aligned}$$

Now, recall that $t^{**}>t^*$, so that if $t>t^{**}$, then $h_A\left(t\right)<h_B\left(t\right)$. Thus, the expression in parenthesis is strictly positive. Thus there is no crossing point to the right of t^{**}. There is also no crossing point to the left of t^{**}, since in that case we could repeat the exercise above to show that t^{**} cannot exist. Thus, t^{**} is unique.

ii) *If there is* t^{**}*so that* $1-G_A\left(t\right)<1-G_B\left(t\right)$ *for* $t<t^{**}$ *and* $1-G_A\left(t\right)>1-G_B\left(t\right)$ *for* $t>t^{**}$*, then* $\Xi_A\left(t\right)>\Xi_B\left(t\right)\ \forall t$.

For $t<t^{**}$ it follows trivially that $\int_0^t G_B\left(s\right)ds<\int_0^t G_A\left(s\right)ds\ \forall t$ with the inequality strict for t above a certain range. For $t>t^{**}$,

$$\frac{\partial\int_0^t\left[G_B\left(s\right)-G_A\left(s\right)\right]ds}{\partial t}=G_B\left(t\right)-G_A\left(t\right)>0.$$

This means that we can bound $\int_0^t [G_B(s) - G_A(s)]\, ds$ above as follows:

$$\int_0^t [G_B(s) - G_A(s)]\, ds < \int_0^\infty [G_B(s) - G_A(s)]\, ds \; \forall t \geq t^{**}$$

Using integration by parts plus the condition that the expected values are the same implies that the bound is zero:

$$\int_0^\infty [G_B(s) - G_A(s)]\, ds \; = -\left(\Lambda_B^{-1} - \Lambda_A^{-1}\right) = 0$$

Thus $\int_0^t G_B(s)ds < \int_0^t G_A(s)ds \; \forall t$. We can verify from equation (8) that $\omega(t) = \Lambda \int_0^t (1 - G(s))\, ds$. Hence, it follows that $1 - \omega_B(s) > 1 - \omega_A(s) \; \forall t$. Since, from equation (9), $1 - \omega(t) = e^{-\Lambda t - \Lambda \Xi(t)}$, it follows that $\Xi_A(t) > \Xi_B(t) \; \forall t$.

2) Suppose the two functions do not cross. The either $h_A(t) > h_B(t)$ for all t or vice versa. In the first case, we have that $G_A(t) = 1 - e^{-\int_0^t h_A(v)dv} > 1 - e^{-\int_0^t h_B(v)dv} = G_B(t)$ for all t (and vice versa in the opposite case). But both of these violate the condition that $\int_0^\infty (1 - G_A(t))\, dt = \int_0^\infty (1 - G_B(t))\, dt$ which is necessary for $\Lambda_A^{-1} = \Lambda_B^{-1}$.

Let t^* be a crossing point. Then, for any $t < t^*$, $h_A(t) = h_A(t^*) - \int_t^{t^*} \frac{\partial h_A(s)}{\partial s} ds$ and $h_B(t) = h_B(t^*) - \int_t^{t^*} \frac{\partial h_B(s)}{\partial s} ds$. Since $\frac{\partial h_A(s)}{\partial s} < \frac{\partial h_B(s)}{\partial s}$ and $h_A(t^*) = h_B(t^*)$ it follows that $h_A(t) > h_B(t)$. Likewise, for any $t > t^*$, $h_A(t) = h_A(t^*) + \int_t^{t^*} \frac{\partial h_A(s)}{\partial s} ds$ and $h_B(t) = h_B(t^*) + \int_t^{t^*} \frac{\partial h_B(s)}{\partial s} ds$ so that $h_B(t) > h_A(t)$. Thus there is a single crossing point and part 1) applies. □

Proof of Proposition 2'. The proposition relies on the fact that we can build a one-sector economy characterized by $\tilde{G}(t) = E\left[\frac{\Lambda_k}{E[\Lambda_k]} G_k(t)\right]$ and $\tilde{\Lambda} = E[\Lambda_k]$ in which monetary shocks have the same real effects as in the heterogeneous economy.[15] Given $\alpha = 1$, firms in each sector do not interact with one another, so that each sector behaves as if it were a separate economy. From **Proposition 1** the real impact of the shock in an economy characterized by G_k is $\Gamma_k = \int_0^\infty e^{-\rho t} (1 - \omega_k(t)) \left(m^{new}(t) - m^{old}\right) dt$, where $\omega_k(t) = \Lambda_k \int_0^t (1 - G_k(t))\, dt$. The real impact of the shock in the multisector economy is just the cross-sectoral average of the real effects in each sector:

$$\Gamma^{het} = E[\Gamma_k] = \int_0^\infty e^{-\rho t} (1 - E[\omega_k(t)]) \left(m^{new}(t) - m^{old}\right) dt.$$

We now show that the real effects of any given monetary shock in the heterogeneous economy described above are identical to the effects in a one sector economy with $\tilde{G}(t) = \frac{E[\Lambda_k G_k(t)]}{E[\Lambda_k]}$. The real impact of the shock in that economy is $\tilde{\Gamma} = \int_0^\infty e^{-\rho t} (1 - \tilde{\omega}(t)) \left(m^{new}(t) - m^{old}\right) dt$, where

[15] Note that this is not the counterfactual one-sector economy discussed subsequently.

$\tilde{\omega}(t) = E[\Lambda_k] \int_0^t \left(1 - \frac{E[\Lambda_k G_k(t)]}{E[\Lambda_k]}\right) dt$. Since

$$E[\omega_k(t)] = E\left[\Lambda_k \int_0^t (1 - G_k(s))\, ds\right] = E[\Lambda_k] \int_0^t \left(1 - \frac{E[\Lambda_k G_k(s)]}{E[\Lambda_k]} ds\right) ds = \tilde{\omega}(t),$$

it follows that $\tilde{\Gamma} = \Gamma^{het}$.

Finally, we can show that the average frequency of price changes in this one-sector economy is the same as in the multisector economy. It is $\tilde{\Lambda} = \left[\int_0^\infty \frac{E[\Lambda_k(1-G_k(t))]}{E[\Lambda_k]} dt\right]^{-1} = \left[\frac{E\left[\Lambda_k \int_0^\infty (1-G_k(t))dt\right]}{E[\Lambda_k]}\right]^{-1} = \left[\frac{E\left[\Lambda_k \Lambda_k^{-1}\right]}{E[\Lambda_k]}\right]^{-1} = E[\Lambda_k]$.

Note that $\tilde{G}(t) = G^{het}(t)$ defined in the text, just as $\tilde{\omega}(t) = \omega^{het}(t)$. Thus, $\mu^{het}(t)$ and $\Xi^{het}(t)$ correspond to selection and cumulative selection for the one-sector economy with the same real effect and the same average frequency of price changes as the multisector economy. The last step of the proof then follows by applying Proposition 2. □

Proof of Proposition 4. First, we note that for the counterfactual one-sector economy,

$$1 - \omega^{count}(t) = e^{-E[\Lambda_k]t - \Xi^{count}(t)},$$

where $\omega^{count}(t) \equiv \int_0^t \left(1 - G^{count}(s)\right) ds$ and for the heterogeneous economy,

$$1 - \omega^{het}(t) = e^{-E[\Lambda_k]t - \Xi^{het}(t)},$$

so that $\Xi^{count}(t) \geq \Xi^{het}(t)$ for all t so long as $1 - \omega^{count}(t) \leq 1 - \omega^{het}(t)$ for all t.

Use **Definition 4** and integrate both sides of equation (8) to write $1 - \omega^{count}(t)$:

$$\begin{aligned} 1 - \omega^{count}(t) &= E[\Lambda_k] \int_t^\infty \left(1 - \bar{G}(E[\Lambda_k]s)\right) ds, \\ &= E[\Lambda_k] \int_{E[\Lambda_k]t}^\infty \left(1 - \bar{G}(v)\right) E[\Lambda_k]^{-1} dv, \\ &= \int_{E[\Lambda_k]t}^\infty \left(1 - \bar{G}(v)\right) dv, \\ &= 1 - \bar{\omega}(E[\Lambda_k]t). \end{aligned}$$

We can show that any $\omega(t) = \Lambda \int_0^t (1 - G(s))ds$ is concave, i.e.:

$$\omega\left(\lambda x + (1-\lambda) x'\right) > \lambda \omega(x) + (1-\lambda)\omega\left(x'\right).$$

This is trivial to see if $G(t)$ is differentiable, since in that case $\frac{\partial^2 \omega(t)}{(\partial t)^2} = -\Lambda \frac{\partial G(t)}{\partial t} < 0$. More generally, given its definition, ω is concave if and only if

$$\int_0^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds > \lambda \int_0^{x} [1-G(s)]\,ds + (1-\lambda)\int_0^{x'} [1-G(s)]\,ds.$$

W.l.o.g., let $x' > x$. Then, we can write

$$\begin{aligned}
& \int_0^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds > \lambda \int_0^{x} [1-G(s)]\,ds + (1-\lambda)\int_0^{x'} [1-G(s)]\,ds \\
\iff\ & \lambda \left[\begin{array}{c} \int_0^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds \\ -\int_0^{x} [1-G(s)]\,ds \end{array}\right] > (1-\lambda)\left[\begin{array}{c} \int_0^{x'} [1-G(s)]\,ds \\ -\int_0^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds \end{array}\right] \\
\iff\ & \lambda \int_x^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds > (1-\lambda)\int_{\lambda x+(1-\lambda)x'}^{x'} [1-G(s)]\,ds \\
\iff\ & (1-\lambda)\left(x'-x\right)\lambda \frac{\int_x^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds}{(1-\lambda)(x'-x)} > \lambda(1-\lambda)\left(x'-x\right)\frac{\int_{\lambda x+(1-\lambda)x'}^{x'} [1-G(s)]\,ds}{\lambda(x'-x)} \\
\iff\ & \frac{\int_x^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds}{(1-\lambda)(x'-x)} > \frac{\int_{\lambda x+(1-\lambda)x'}^{x'} [1-G(s)]\,ds}{\lambda(x'-x)}.
\end{aligned}$$

We can verify that the inequality holds since G is increasing, implying that

$$\frac{\int_x^{\lambda x+(1-\lambda)x'} [1-G(s)]\,ds}{(1-\lambda)(x'-x)} > \frac{\int_x^{\lambda x+(1-\lambda)x'} [1-G(\lambda x+(1-\lambda)x')]\,ds}{(1-\lambda)(x'-x)} = 1-G\left(\lambda x+(1-\lambda)x'\right),$$

and

$$\frac{\int_{\lambda x+(1-\lambda)x'}^{x'} [1-G(s)]\,ds}{\lambda(x'-x)} < \frac{\int_{\lambda x+(1-\lambda)x'}^{x'} [1-G(\lambda x+(1-\lambda)x')]\,ds}{\lambda(x'-x)} = 1-G\left(\lambda x+(1-\lambda)x'\right)$$

In particular, given that any ω is concave, so is $\bar{\omega}(t) \equiv 1-\int_0^t \left(1-\bar{G}(s)\right)ds$ with $\bar{G}(s)$ as defined in the statement of the proposition is also concave. Therefore, from Jensen's inequality, it follows that $E[\bar{\omega}(\Lambda_k t)] < \bar{\omega}(E[\Lambda_k t])$, and

$$1-\omega^{het}(t) = 1-E[\bar{\omega}(\Lambda_k t)] > 1-\bar{\omega}(E[\Lambda_k]\,t) = 1-\omega^{count}(t).$$

The last step of the proof follows from $1-\omega(t) = e^{-\Lambda-\Lambda\Xi(t)}$. Since, by construction, $\Lambda^{het} = \Lambda^{count} = E[\Lambda_k]$, and since $1-\omega^{het}(t) > 1-\omega^{count}(t)$, it follows that $\Xi^{het}(t) < \Xi^{count}(t)$.

□

Proposition 6 is a special case of Proposition 6', proved below:

Proof of Proposition 6'. Consider first a case with bounded support, i.e., there is z such that $\omega(t) = 1 \ \forall t \geq z$:

$$\begin{aligned}
\lim_{\rho\to 0}\Gamma &= \int_0^\infty \sum_{k=1}^K (1-\omega(t))\, a_k t^{k-1} dt = \\
&= \int_0^z \sum_{k=1}^K (1-\omega(t))\, a_k t^{k-1} dt = \\
&= \sum_{k=1}^K a_k \left[\frac{z^k}{k}(1-\omega(z)) - 0\times(1-\omega(0)) - \int_0^z \frac{t^k}{k}\frac{\partial(1-\omega(t))}{\partial t}dt\right] = \\
&= 0 + \sum_{k=1}^K a_k \int_0^z \frac{t^k}{k}\frac{\partial\omega(t)}{\partial t}dt = \\
&= \Lambda \int_0^z \sum_{k=1}^K a_k \frac{t^k}{k}(1-G(t))\, dt.
\end{aligned}$$

Then,

$$\begin{aligned}
\lim_{\rho\to 0}\Gamma &= \Lambda \int_0^z \sum_{k=1}^K a_k \frac{t^k}{k}(1-G(t))\, tdt = \\
&= \Lambda\left[\sum_{k=1}^K a_k \frac{z^{k+1}}{k(k+1)}(1-G(z)) - 0\times(1-G(0)) - \int_0^z \frac{t^{k+1}}{k(k+1)}d(1-G(t))\right] = \\
&= \Lambda \int_0^z \sum_{k=1}^K a_k \frac{t^{k+1}}{k(k+1)}dG(t) = \\
&= \Lambda \sum_{k=1}^K \frac{a_k}{k(k+1)} E\left[\tau^{k+1}\right] = \\
&= \sum_{k=1}^K \frac{a_k}{k(k+1)}\frac{E\left[\tau^{k+1}\right]}{E[\tau]},
\end{aligned}$$

where the last line follows from $\Lambda^{-1} = E[\tau]$.

The case with unbounded support can be obtained by constructing a sequence of distribution functions $G_z(t)$ defined as:

$$G_z(t) = \frac{G(t)}{G(z)}\mathbb{1}(t\leq z) + \mathbb{1}(t\geq z),$$

with associated $\Lambda_z = \left[\int_0^\infty (1 - G_z(t))\, dt\right]^{-1}$. We take the limit

$$\lim_{z\to\infty}\lim_{\rho\to 0}\Gamma = \lim_{z\to\infty}\Lambda_z \int_0^\infty \sum_{k=1}^{K} a_k \frac{t^k}{k}\left(1 - G_z(t)\right) dt.$$

We then use Lebesgue's dominated convergence theorem (see, for example, Kolmogorov and Fomin 1970) to show that this limit is equal to $\Lambda \int_0^\infty a_k \frac{t^k}{k}(1 - G(t))$. Let $\{z_1, z_2, ..., z_n, ...\}$ be an infinite sequence such that $z_{k+1} > z_k\ \forall k$ and $z_1 > 0$. Then, for all t, $\lim_{n\to\infty}\Lambda_{z_n}(1 - G_{z_n}(t))\, t^k = \Lambda(1 - G(t))\, t^k$. Furthermore there is a $\bar{\Lambda} < \infty$ such that, $\Lambda_{z_n}(1 - G_{z_n}(t))\, t^k < \bar{\Lambda}(1 - G(t))\, t^k$. That $1 - G_{z_n}(t) < 1 - G(t)$ follows trivially from the definition of G_{z_n}. To see that such a $\bar{\Lambda}$ exists, recall that $\Lambda_{z_n} = \left[\int_0^\infty (1 - G_{z_n}(t))\, dt\right]^{-1}$ and that for all $n > 1$, $\int_0^\infty (1 - G_{z_n}(t))\, dt = \int_0^{z_n}\left(1 - \frac{G(t)}{G(z_n)}\right) dt > \int_0^{z_1}\left(1 - \frac{G(t)}{G(z_1)}\right) dt$. Thus, it is enough to pick $\bar{\Lambda} > \left[\int_0^{z_1}\left(1 - \frac{G(t)}{G(z_1)}\right) dt\right]^{-1}$.

If $\int_0^\infty (1 - G(t))\, t^k dt < \infty$ then $\bar{\Lambda}\int_0^\infty (1 - G(t))\, t^k dt < \infty$ and all the conditions of the theorem are satisfied. It follows that $\lim_{n\to\infty}\int_0^\infty (1 - G_{z_n}(t))\, t^k dt = \int_0^\infty (1 - G(t))\, t^k dt$ and $\Lambda\int_0^\infty \sum_{k=1}^{K} a_k \frac{t^k}{k}(1 - G(t))\, dt = \lim_{z\to\infty}\Lambda_z \int_0^\infty \sum_{k=1}^{K} a_k \frac{t^k}{k}(1 - G_z(t))\, dt$. □

A.4 Implementation of the Numerical Simulation

The numerical analysis is based on a log-linearized, discrete-time version of the model solved using Dynare. The heart of the model is a pricing rule, dependent on α. Let x_t be the reset price chosen by firms who get to choose their prices at t in log-deviation form. The discrete time analogue of equation (4) is

$$x_t = \Lambda \sum_{s=0}^{\infty} \beta^s (1 - G_s)\left[\alpha m_{t+s} + (1 - \alpha)\, p_{t+s}\right] ds, \tag{30}$$

The price level at t is:

$$p_t = \Lambda \sum_{s=0}^{\infty} (1 - G_s)\, x_{t-s}$$

and output is

$$y_t = m_t - p_t$$

The law of motion for m_t is:

$$\Delta m_t = 0.5\Delta m_{t-1} + \varepsilon_t$$

Let y^{IRF}_{t+s} be the impulse response function of a unit shock to ε_t that hits the economy at t. We calculate the cumulative real effect as

$$\Gamma = \sum_{s=0}^{\infty} \beta^s y^{IRF}_{t+s}$$

This system of four equations in four variables describes how output reacts to shocks. Note that m_t is not stationary, so, as written this model is not amenable to be solved using conventional methods. We can make the model stationary by rewriting it in terms of $\tilde{p}_t = p_t - m_t$ and $\tilde{x}_t = x_t - m_t$. Then we have the equivalent model:

$$\tilde{x}_t = \Lambda \sum_{s=0}^{\infty} \beta^s (1 - G_s) \left[\sum_{v=0}^{s} \Delta m_{t+v} + (1-\alpha)\tilde{p}_{t+s} \right] ds, \tag{31}$$

$$\tilde{p}_t = \Lambda \sum_{s=0}^{\infty} (1 - G_s) \left(\tilde{x}_{t-s} - \sum_{v=0}^{s} \Delta m_{t-s+v} \right) \tag{32}$$

$$y_t = -\tilde{p}_t \tag{33}$$

$$\Delta m_t = 0.5\Delta m_{t-1} + \varepsilon_t \tag{34}$$

This model as written is still not solvable in Dynare because it involves infinite sums, even if convergent. We solve it for cases where there is some J such that $G_s = 1$ for all $s > J$. In those cases the state space becomes finite.

For the truncated exponentials case, we have that, given the maximum duration T and the decay parameter θ, the average duration of price-spells is:

$$E[\tau] = \left(1 - (1-\theta)^{T-1}\right)\frac{1}{\theta} + (1-\theta)^{T-1}$$

For any T, we look for θ such that $E[\tau] = 2$. We consider $T = \{2, ..., 20\}$. Table (1) gives the corresponding θ's, together with the mean, variance and skewness of price durations for some of these cases. Note that the variance is increasing in T.

Table 1: Parameters and Moments for Truncated Exponential Models

T	θ	mean	variance	skewness
2	0.000	2.000	0.000	-
4	0.456	2.000	1.234	0.703
6	0.491	2.000	1.718	1.430
8	0.498	2.000	1.903	1.804
10	0.500	2.000	1.969	1.988
12	0.500	2.000	1.990	2.070
14	0.500	2.000	1.997	2.103
16	0.500	2.000	1.999	2.115