

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

Carlos Carvalho  
PUC-Rio

Felipe Schwartzman  
Federal Reserve Bank of Richmond

December 2012

## Abstract

Given the frequency of price changes, the real effects of a monetary shock are smaller if adjusting firms are disproportionately likely to be ones with prices set before the shock. This *selection effect* is important in a large class of sticky-price models with time-dependent price adjustment. We provide a very general analytical characterization of the relationship between this selection effect, the distribution of the duration of price spells, and the real effects of monetary shocks. We find that: 1) Selection is stronger and real effects are smaller if the hazard function of price adjustment is more strongly increasing; 2) Selection is weaker and real effects are larger if there is sectoral heterogeneity in price stickiness; 3) Selection is stronger and real effects are smaller if the durations of price spells are less variable. We also show that 4) If monetary shocks affect primarily the level of nominal aggregate demand, the mean and variance of price durations are sufficient statistics for the real effects of such shocks.

JEL classification codes: E1, E3

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

---

\*For comments and suggestions, we thank participants at LAMES/LACEA 2012, EEA 2012, SED 2012, ESWM 2012, and ESEM 2011. 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 micro level do not automatically imply that monetary disturbances have large real effects at the macro level. For the same frequency of price changes, the real effects of a monetary shock are small if adjusting firms are also the ones that are most likely to react strongly to the shock. The importance of this *selection effect* has been well understood at least since Caplin and Spulber (1987). In their model, a small fraction of firms making large price adjustments is enough to completely offset monetary shocks and induce money neutrality.

In Caplin and Spulber (1987) – and in menu-cost models more generally – the selection effect that dampens monetary non-neutralities arises from *self-selection*: firms always have the option of incurring a menu cost to adjust their prices,<sup>1</sup> so that not only prices, but the fraction of firms adjusting prices reacts to shocks. Furthermore, there is a selection effect because adjusting firms are also the ones for which the current price is far enough from the target to justify paying the menu cost.

We argue that selection effects do not necessarily hinge on self-selection. In particular, selection effects are important in environments where the probability of a firm adjusting a price only depends on the time elapsed since the price was put in place and where the proportion of firms adjusting prices at any given time is constant. This is important because such time-dependent pricing rules are prevalent in the sticky-price literature.

In this paper we study selection and monetary non-neutrality in a large class of time-dependent sticky-price models. As in Dotsey, King, and Wolman (1997, section 3), price changes arrive according to a general function of the time elapsed since the last price change. At any point in time after a monetary shock, firms are partitioned into two groups: those who have prices that pre-date the shock (henceforth “old prices”), and those whose price post-dates the shock (henceforth “new prices”). We say that there is *positive selection* when old prices are over-represented among adjusting prices and *negative selection* otherwise. The constant hazard of adjustment model (Calvo, 1983) is the knife edge case where there is no selection. The constant duration model (Taylor, 1979) is the extreme case where selection is most positive given the frequency of price changes.

Our goal is to characterize the role of such selection effects in determining the real effects of nominal shocks. We use a scalar measure of the real impact of the shock – the cumulative output gain associated with a one-time unexpected nominal expansion.<sup>2</sup> To isolate the role of selection from other factors that are well known to affect the extent of monetary non-neutrality, we always

---

<sup>1</sup>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.

<sup>2</sup>We show in the Appendix that this measure has a welfare interpretation.

compare economies with the same average frequency of price changes, and, for most of the paper, focus on an environment in which the optimal price for a given firm is neither a strategic substitute nor a strategic complement to the price of other firms – what we refer to as *strategic neutrality in price setting*.

We show that the real effects of a permanent shock to the level of nominal income are larger if selection is smaller. Furthermore, we show that, for a given average frequency of price changes, real effects depend solely on selection, and therefore the latter contains all the relevant information about price-setting frictions. It follows that if two economies with the same average frequency of price changes produce different real effects in reaction to nominal shocks, then it must be the case that price changing firms are selected differently.

We find that selection is a useful organizing principle for interpreting previous findings in the literature:

1) Calvo (1983) contracts imply no selection, as the probability of price changes does not depend on the time elapsed since the last price change (henceforth the “age” of the price). In contrast, Taylor (1979) contracts imply the highest possible 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 (1979) pricing implies smaller real effects than Calvo (1983) pricing (Kiley, 2002).

2) If, as argued by Dotsey, King, and Wolman (1997), the hazard of price adjustment is increasing over time, then selection is positive and the real effects of a nominal shock are smaller than under Calvo contracts.<sup>3</sup> We also show that the more increasing the hazard of price adjustment is, the larger selection effects are. These findings follow from the simple intuition that a more increasing hazard function implies that older prices have a relatively higher probability of being changed.<sup>4</sup>

3) Heterogeneity in price stickiness is associated with more negative selection, as sectors with low frequency of price changes are both more likely to have a large proportion of “old” prices and to have a lower probability of having their firms selected for price changes. This clarifies and generalizes the finding in Carvalho (2006) that the real effects of a nominal shock are larger in an economy with heterogeneity in price stickiness than in an otherwise identical economy with the same average frequency of price changes.

4) Relatedly, when comparing two economies, one in which the distribution of the duration of price spells is a mean preserving spread of the other, then selection is lower in the economy with the more variable price spells, and the real effects of the shock are larger.

---

<sup>3</sup>Wolman (1999) provides a similar intuition.

<sup>4</sup>Sheedy (2010) also explores models with increasing hazard functions and shows that they are able to generate inflation persistence.

5) In the case of permanent level shocks to nominal aggregate demand, the mean and the variance of the duration of price spells are sufficient statistics for the real effects of the shock.<sup>5</sup> This result is of particular interest because it provides guidance to the empirical literature that explores microeconomic price data, and for calibration of models without the need to specify the whole distribution of the duration of price spells.

Lastly, we show how these results generalize. In particular, we show that for any monetary shock that implies a new path for nominal income which is always above the expected one, the real effects of the shock are smaller if selection is higher. Hence, results 1-4 continue to hold in this more general case. In terms of result 5, we also show that if a shock affects the growth rate (as opposed to the level) of nominal income, then the third moment of the distribution of price spells also matters. Finally, we run numerical simulations where we show that the average and the variance of price spells is close to being sufficient statistics to describe the contribution of the distribution of the duration of price spells to the real effects of empirically plausible nominal shocks. We find that this is true not only under strategic neutrality, but also under strategic complementarity and strategic substitutability in price setting.

Our paper is not the first one to identify a role for selection in time-dependent pricing models. 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.

## 2 Model

A representative household derives utility from a continuum of differentiated consumption goods aggregated in a Dixit-Stiglitz composite, and supplies labor to the monopolistically competitive firms that produce these goods. The household owns these firms, so it receives back whatever profits they generate. Firms hire specific labor types in competitive market. 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. This allows us to resort to the certainty equivalence principle and focus on the dynamic response of the economy to one-time

---

<sup>5</sup>Without linking it to selection, we first proved this result in Carvalho and Schwartzman (2008). In ongoing work, Alvarez et al. (2012) show that it also holds in a model in which the adjustment hazard function arises endogenously from an optimal price-setting problem in the presence of information costs.

shocks in a world of otherwise perfect foresight. The representative household maximizes:

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

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

and subject to a no-Ponzi condition. Here  $\rho$  is the discount rate,  $C(t)$  is consumption of the composite good,  $L_j(t)$  is the quantity of labor supplied for the production of variety  $j$ ,  $W_j(t)$  is the nominal wage for labor in variety  $j$ ,  $T(t)$  are firms' flow profits received by the consumer,  $B(t)$  denotes bond holdings that accrue interest at rate  $i(t)$ , and  $P(t)$  is a price index to be defined below.

The composite consumption good is given by:

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

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:

$$\frac{W_j(t)}{P(t)} = C(t)^\sigma L_j(t)^{\frac{1}{\psi}}, \quad (2)$$

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

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

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 price-setting models include Taylor (1979) and Calvo (1983). We follow Dotsey, King, and Wolman (1997), and generalize these standard frameworks. We denote the probability of a price surviving for less than  $s$  by a generic

cumulative distribution function  $G(s)$ . The only restrictions we impose are that a)  $G(s)$  depends only on the time elapsed since the price was last reset, but not on the particular date in which it was reset, b) that  $\lim_{s \rightarrow \infty} G(s) = 1$  so that all price spells come to an end with probability one. Certain results will also require restrictions on  $G$  that we will introduce as needed.

A firm that sets a new price at time  $t$  chooses the price  $X_j(t)$  to solve:

$$\begin{aligned} \max_{X_j(t)} \int_0^\infty Q(t, t+s) (1 - G(s)) [X_j(t) Y_j(t+s) - W_j(t+s) N_j(t+s)] ds \\ \text{s.t. } Y_j(t+s) = N_j(t+s), \\ Y_j(t+s) = \left( \frac{X_j(t)}{P(t+s)} \right)^{-\varepsilon} Y(t+s), \end{aligned}$$

where  $Q(t, t+s)$  is a discount factor,  $N_j(t+s)$  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:

$$X_j(t) = \frac{\varepsilon}{\varepsilon - 1} \frac{\int_0^\infty Q(t, t+s) (1 - G(s)) P(s)^\varepsilon Y(s) W_j(s) ds}{\int_0^\infty Q(t, t+s) (1 - G(s)) P(s)^\varepsilon Y(s) ds}.$$

We focus on the symmetric equilibrium in which, conditional on time- $t$  information, the joint distribution of future variables that matter for price setting is the same for all firms, and therefore they choose the same nominal price  $X_j(t)$ . This allows us to drop the  $j$ 's and denote the price set by any firm at time  $t$  as  $X(t)$ . We moreover assume uniform staggering of pricing decisions, so that the price index satisfies:

$$P(t) = \left[ \int_{-\infty}^t \Lambda (1 - G(t-v)) X(v)^{1-\varepsilon} dv \right]^{\frac{1}{1-\varepsilon}}, \quad (3)$$

where  $\Lambda^{-1} \equiv \int_0^\infty (1 - G(s)) ds$ . 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. We thus refer to  $\Lambda$  as the average frequency of price changes in the economy.

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 policy unspecified and assume an exogenous path for nominal aggregate demand,  $M(t) = P(t) Y(t)$ .

We log-linearize the model around a deterministic, zero-inflation symmetric steady state. In

this log-linear environment, firms that change prices at time  $t$  set (lowercase variables denote log-deviations from the steady state):

$$x(t) = \frac{\int_0^\infty e^{-\rho s} (1 - G(s)) w_j(t+s) ds}{\int_0^\infty e^{-\rho s} (1 - G(s)) ds}. \quad (4)$$

Log-linearizing the first-order condition for consumption/leisure choice of the household yields:

$$w_j(t+s) = p(t+s) + \sigma c(t+s) + \psi^{-1} l_j(t+s).$$

From the production function, we have  $l_j(t+s) = y_j(t+s)$  and from market clearing,  $c(t+s) = y(t+s)$ . Moreover, the demand function for varieties in log-linear form tells us that  $y_j(t+s) = y(t+s) - \varepsilon(x(t) - p(t+s))$ , where  $x(t)$  is the price currently chosen by the firm. Bringing all of these together we can express the wage rate  $w_j(t+s)$  in terms of aggregates and  $x(t)$ :

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

We can also use  $m(t+s) = p(t+s) + y(t+s)$  to substitute out  $y(t+s)$  and obtain:

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

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

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

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 is a knife-edge case that requires specific constellations of primitive parameters such as  $\sigma = 1$  and  $\psi \rightarrow \infty$  (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

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

### 3 The real effects of a monetary shock

Throughout most of the paper, we examine the commonly used case of a shock to the level of nominal aggregate demand in an economy with strategic neutrality in price setting. This allows us to isolate the role of selection from the well-known effects of interactions between firms' pricing decisions. In the final sections we provide an analytical discussion of the effect of more general shocks and show with simulations that the main results survive the presence of pricing interactions.

The economy starts with a constant level of nominal aggregate demand  $M_0$ , with associated pricing decisions  $X^{old}$ , the aggregate price level  $P^{old}$ , and constant output  $Y^{old}$ . We then 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 permanent level for nominal aggregate demand  $M^{new} = M_0 e^{\Delta m}$ , and associated paths for pricing decisions, aggregate price level, and output – respectively,  $X^{new}(t)$ ,  $P^{new}(t)$ , and  $Y^{new}(t)$ . Because we focus on a log-linear approximation of the model around the deterministic zero-inflation steady state, in which certainty equivalence holds, the assumption of a one-time unforeseen shock allows us to derive the impulse response functions of endogenous variables to the shock while simplifying the exposition considerably. 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.

$$m^{old}(t) = m_0, \quad m^{new}(t) = m_0 + \Delta m. \quad (6)$$

Firms choose different prices depending on whether their choice occurs before or after the shock hits. Due to the absence of pricing interactions ( $\alpha = 1$ ), from (5) it follows immediately that

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

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

$$p^{new}(t) = p(t) = \omega(t) m^{new} + (1 - \omega(t)) m^{old}, \quad (7)$$

where  $\omega(t) \equiv \Lambda \int_0^t 1 - G(v) dv$  is the fraction of firms with *new* prices (i.e., who last set their prices after the shock).

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

$$y^{new}(t) - y^{old} = m^{new} - p^{new}(t) - (m^{old} - p^{old}) = \Delta m (1 - \omega(t)). \quad (8)$$



In words, the output effect at  $t$  is proportional to the size of the shock  $\Delta m$  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 at  $t$  is larger.

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:<sup>6</sup>

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

We refer to  $\Gamma$  generically as *the real effects of the monetary shock*.

Using (8), we can write the real effects of the shock as:

$$\Gamma = \Delta m \int_0^{\infty} e^{-\rho t} (1 - \omega(t)) dt. \quad (9)$$

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.

## 4 Selection and monetary non-neutrality

In this section we introduce the concept of selection that we analyze in this paper, and how we measure it. In statistics, we say there is 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, our sample consists of the prices being reset at a given point in time, and the population encompasses all existing prices. The sample moment we focus on is the fraction of old prices being reset at  $t$ , and the population moment is the fraction of old prices still in place at  $t$ .

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

**Definition 1** *Selection* (at  $t$ ), denoted by  $\mu(t)$ , is defined as

$$\mu(t) \equiv \frac{1 - G(t)}{1 - \omega(t)} - 1. \quad (10)$$

---

<sup>6</sup>We show in the Appendix that, up to a first-order approximation, this measure is proportional to the ex-post utility impact generated by the shock in an economy with a distorted steady state.

A related measure of selection emphasizes not the selection at a given point in time, but the 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 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 the converse is not necessarily true.

We are now ready to describe the role of selection in determining the extent of monetary non-neutrality. Note that, 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,  $\Lambda$ ) and  $b)$  are doing so for the first time after the shock ( $1 - G(t)$ ):

$$\frac{\partial \omega(t)}{\partial t} = \Lambda(1 - G(t)). \quad (11)$$

Solving the differential equation (11) and using the definitions above yields:

$$1 - \omega(t) = e^{-\Lambda \int_0^t \frac{1-G(v)}{1-\omega(v)} dv} = e^{-\Lambda t - \Lambda \int_0^t \mu(v) dv} = e^{-\Lambda t - \Lambda \Xi(t)}. \quad (12)$$

It follows from the definition of the real effects of the monetary shock (equation 9) and from the solution to the differential equation (11) that

$$\frac{\Gamma}{\Delta m} = \int_0^\infty e^{-\rho t} (1 - \omega(t)) dt = \int_0^\infty e^{-(\rho+\Lambda)t - \Lambda \int_0^t \mu(v) dv} dt = \int_0^\infty e^{-(\rho+\Lambda)t - \Lambda \Xi(t)} dt. \quad (13)$$

We can thus derive three important immediate implications:<sup>7</sup>

**Proposition 1** *Given  $\Lambda$ ,*

- 1) *The real effects of a monetary shock are larger if selection,  $\mu(t)$ , is smaller for all  $t$ .*
- 2) *The real effects of a monetary shock are larger if cumulative selection,  $\Xi(t)$ , is smaller for all  $t$ .*

3) *Let  $\mathcal{M}$  be the set of all functions  $\mu : [0, \infty) \rightarrow \mathbb{R}$  that can be constructed using **Definition 10**. Let  $\mathcal{G}^\Lambda$  be the set of all functions  $G : [0, \infty) \rightarrow [0, 1]$  satisfying  $\int_0^\infty (1 - G(s)) ds = \Lambda^{-1}$ . Then, there is a mapping  $f : \mathcal{M} \rightarrow \mathcal{G}^\Lambda$  that allows us to recover  $G$  from  $\mu$ .*

The last part of the proposition states formally that, given  $\Lambda$ ,  $G$  can be obtained from  $\mu$ , so that given  $\Lambda$ ,  $\mu$  and  $G$  are equally valid primitives for the general class of time-dependent pricing

---

<sup>7</sup>All proofs are in the Appendix.

models that we consider. This is important because it means that, once we know  $\Lambda$ , any interesting properties of the c.d.f. of price durations  $G$  can be expressed in terms of properties of *selection*  $\mu$ .

**Example 1: Calvo pricing**

A leading example of time-dependent pricing used in the literature is the one proposed by Calvo (1983). In his model, 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 changed its price. This implies an exponential decay of the survival probability of a price. In terms of our notation:

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

Also, it is easy to verify that

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

so that selection equals

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

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

The real effects of a permanent shock to the level of nominal aggregate demand under Calvo pricing for a small  $\rho$  are:

$$\lim_{\rho \rightarrow 0} \Gamma = \Delta m \int_0^\infty e^{-\Lambda t} dt = \frac{\Delta m}{\Lambda}. \quad \blacksquare$$

**Example 2: Taylor pricing**

A second leading example is based on Taylor (1979). In this model firms set prices for a fixed period of time (given by  $\Lambda^{-1}$ ). Thus, the distribution of price durations is degenerate at  $\Lambda^{-1}$ .

In terms of our notation:

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

so that the selection is given by:

$$\mu(t) = \begin{cases} \frac{1}{1-\Lambda t} - 1 & \text{if } t < \Lambda^{-1} \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Selection is positive wherever it is defined (i.e., wherever the pool of old prices is not empty). In fact, within that range, selection is maximal, since all price changes come from the pool of old

prices. To show this formally, note that under Taylor pricing, for any  $t < \Lambda^{-1}$ ,

$$\mu(t) = \frac{1}{1 - \Lambda t} - 1,$$

whereas under any alternative pricing rule with the same average frequency of price changes  $\Lambda$ ,

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

That  $1 - G(t) \leq 1$  is immediate. We can also show that  $1 - \omega(t) \geq 1 - \Lambda t$ :

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

Cumulative selection is:

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

Note, however, that  $\lim_{t \rightarrow \Lambda^{-1}} \mu(t) = \infty$ . We can extend the definition of  $\Xi(t)$  to  $t > \Lambda^{-1}$  as follows:

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

Since under Taylor pricing selection is the highest possible for a given  $\Lambda$ , it follows that cumulative selection is the highest possible. Thus, for a given  $\Lambda$ , Taylor pricing implies the smallest real effects among all time-dependent pricing models.<sup>8</sup>

The real effects of a shock under Taylor pricing for a small  $\rho$  are:

$$\begin{aligned} \lim_{\rho \rightarrow 0} \Gamma &= \Delta m \int_0^{\infty} \min\{1 - \Lambda t, 0\} dt \\ &= \Delta m \int_0^{\Lambda^{-1}} (1 - \Lambda t) dt \\ &= \frac{1}{2} \frac{\Delta m}{\Lambda}, \end{aligned}$$

which amounts to half the real effects obtained under Calvo pricing with the same average frequency of price changes. ■

The result that Calvo pricing implies larger real effects than Taylor pricing is not new (Kiley, 2002). What we show is that the difference is due to the strong selection under Taylor pricing and the complete lack of selection under Calvo pricing.

We now turn to general results concerning how the shape of the distribution of the duration of

---

<sup>8</sup>Without linking it to selection, we first proved this result in Carvalho and Schwartzman (2008) – a working paper that we never submitted for publication.

price spells affects the real effects of monetary shocks. We show how these different results relate to the overarching notion of selection.

#### 4.1 Hazard functions

The empirical literature on price-setting has devoted quite a bit of effort to measuring the shape of the hazard function of price adjustment.<sup>9</sup> The hazard function is defined as:

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

Both Calvo (1983) and Taylor (1979) make strong assumptions about the shape of the hazard function and, at least since the work of Dotsey, King, and Wolman (1997) and Wolman (1999) it has been clear that the shape of the hazard function matters for the real effects of monetary shocks.

We start by showing that the concept of selection is closely related to that of the hazard function. Specifically, the following is true:

**Lemma 1** *Let  $\mu$  and  $h$  be, respectively, the selection function and hazard function associated with a c.d.f.  $G$ . Then,*

$$\mu(t) = \int_t^\infty \frac{h(s)}{\Lambda} \Psi_t(s) ds - 1, \tag{14}$$

where

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

The intuition for the result above is as follows. The hazard function is the continuous-time analogue of the probability that a price will change at a given age, conditional on it having survived up to that age. Selection is up to a constant, proportional to the continuous-time analogue of the probability that a price will change at a given age conditional on it having survived up to that age *or longer*. Thus, selection is related to the appropriately weighted average of hazard functions of prices older than  $t$ . More specifically, selection at  $t$  is, up to a constant, proportional to the average of the hazard functions across all prices that are older than  $t$  in the population. The constant of proportionality is the average duration of price spells,  $\Lambda^{-1}$ , and the weights  $\Psi_t(s)$  are the fractions of prices of age  $s$  among all prices older than  $t$  in the population.

From (14), we can show that if a hazard function is (everywhere) increasing – as suggested by Dotsey, King, and Wolman (1999) – then there is positive selection (everywhere):

---

<sup>9</sup>Klenow and Malin (2010, section 5.3) provide a review of the empirical literature estimating hazard functions.

**Lemma 2** 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 that the probability of a price change increases with the age of the price. Thus, older prices are over-selected relative to the case of a flat hazard function, in which there is no selection.

The general intuition extends to the comparison of two hazard functions. In this case, the connection is with cumulative selection. Given two economies, one with a more increasing hazard function than the other, we can show that the economy with the more increasing hazard function will feature higher cumulative selection and lower monetary non-neutrality:<sup>10</sup>

**Proposition 2** 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:

1) there is a  $t^*$  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  $\Xi_A(t) < \Xi_B(t) \forall t$ .

An immediate corollary follows:

**Corollary 1** The real effects of a level shock are larger in economy  $A$  than in  $B$ :  $\Gamma_A > \Gamma_B$ .

### Example 1: Calvo pricing

With Calvo pricing, the c.d.f. of durations is given by  $1 - e^{-\Lambda t}$ . The hazard function is:

$$h(t) = \frac{\Lambda e^{-\Lambda t}}{e^{-\Lambda t}} = \Lambda,$$

so that Calvo pricing implies a constant hazard function. Selection is:

$$\mu(t) = \int_t^\infty \frac{h(s)}{\Lambda} \Psi_t(s) ds - 1 = \int_t^\infty \frac{\Lambda}{\Lambda} \Psi_t(s) ds - 1 = 0.$$

The fact that Calvo pricing implies a constant hazard function for price changes provides some additional intuition for why selection is zero in this case. A constant hazard function means that the probability of a price change does not depend on the age of the price. Therefore, prices that are being changed are, by necessity, representative of the age distribution of the population.

<sup>10</sup>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).

One implication of **Lemma 2** is that, for a given frequency of price changes, Calvo pricing generates the largest real effects among economies with non-decreasing hazard functions. ■

### Example 2: Taylor pricing

We return to the constant duration model proposed by Taylor (1979). In this setup, firms keep their prices fixed for  $\Lambda^{-1}$  periods and then change it with probability 1. The hazard function is 0 for all  $t \leq \Lambda^{-1}$  and not defined elsewhere. Strictly speaking, **Proposition 2** does not apply to Taylor pricing, since it does not have a well defined hazard function everywhere. Still, we can make a heuristic argument based on an approximation of a discrete time setting that the hazard function under Taylor is more increasing than any alternative. To see this, consider the c.d.f. of the discrete time analogue of the Taylor pricing, given by:

$$G_t = \begin{cases} 0 & \text{if } t < \Lambda^{-1} \\ 1 & \text{if } t \geq \Lambda^{-1}. \end{cases}$$

Suppose each time period has length  $\Delta t$ . The discrete time analogue of the hazard function is:

$$h_t = \frac{\frac{G_t - G_{t-1}}{\Delta t}}{1 - G_{t-1}} = \begin{cases} 0 & \text{for } t < \Lambda^{-1} \\ \frac{1}{\Delta t} & \text{for } t = \Lambda^{-1} \\ \text{undefined} & \text{otherwise.} \end{cases}$$

It follows that, for  $t = \Lambda^{-1}$ ,  $\lim_{\Delta t \rightarrow 0} h_t = \infty$ .

At  $t = \Lambda^{-1}$ ,  $\lim_{\Delta t \rightarrow 0} h_t$  is greater than any alternative hazard function, and at  $t = 0$  the hazard function is simply zero so that it is smaller than any alternative. This implies that, in the limit, there is a single crossing at  $t = \Lambda^{-1}$  with any alternative hazard function, just as required by **Proposition 2**. This provides an alternative proof that the Taylor economy has the smallest real effects among all economies with a given average frequency of price changes. ■

## 4.2 Heterogeneity in price stickiness

We can also appeal to selection to shed light on the findings by Carvalho (2006) that nominal shocks have larger real effects in an economy with heterogeneity in price stickiness than in a one-sector economy with the same frequency of price changes. These findings are of particular importance because, as documented by Bils and Klenow (2004) and subsequent work, there is substantial cross-sectoral heterogeneity in the frequency of price changes.

As a first step, we show that, by choosing a suitable adjustment hazard function, we can represent a heterogeneous economy as a one-sector economy. Consider an economy with  $K$  sectors indexed by  $k$ , each with a measure  $\Phi_k$  of firms. For notational convenience, define cross-sectoral

averages as:

$$E[x_k] \equiv \sum \Phi_k x_k.$$

The c.d.f. of price durations in sector  $k$  is  $G_k(t)$ . The price level in this economy is:

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

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

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

where  $\Lambda_k^{-1} = \int_0^\infty (1 - G_k(s)) ds$ .

Now, consider a permanent shock to the level of nominal aggregate demand as in equation (6). Given this shock, the optimal pricing policy is:

$$\forall k, x_k(t) = \begin{cases} m_0 & \text{if } t \leq 0 \\ m_1 & \text{if } t > 0. \end{cases}$$

Since the pricing policy in all sectors is identical, we can simplify the notation by letting  $x(t) \equiv x_k(t)$ . Then:

$$\begin{aligned} p(t) &= \int_{-\infty}^t E[\Lambda_k (1 - G_k(t-s))] x(s) ds \\ &= \int_{-\infty}^t \Lambda^{het} (1 - G^{het}(t-s)) x(s) ds, \end{aligned}$$

where  $\Lambda^{het} \equiv E[\Lambda_k]$  and  $G^{het}(t) \equiv \frac{E[\Lambda_k G_k(t)]}{E[\Lambda_k]}$ .

Hence, for our purposes, the multi-sector economy can be represented as a one-sector economy with  $\Lambda^{het} = E[\Lambda_k]$  and  $G^{het}(t) = \frac{E[\Lambda_k G_k(t)]}{E[\Lambda_k]}$ . These calculations give rise to the following lemma:

**Lemma 3** *Consider an economy with multiple sectors with a fraction  $\Phi_k$  of firms in sector  $k$  and  $\alpha = 0$  (strategic neutrality in price setting). In this economy, the dynamics of  $p(t)$  and  $y(t)$  in response to a shock to the level of nominal income as described in equation (6) are identical to those of a one-sector economy with c.d.f. of price spells given by  $G^{het}(t)$ , average frequency of price changes  $\Lambda^{het}$ , and fraction of new prices  $\omega^{het}(t)$  such that*

$$G^{het}(t) = \frac{E[\Lambda_k G_k(t)]}{E[\Lambda_k]}, \quad \Lambda^{het} = E[\Lambda_k], \quad \omega^{het}(t) = E[\omega_k(t)],$$

where the operator  $E$  takes the cross-sectional average given the weights  $\Phi_k$ .

The observational equivalence result relies on the fact that, for this type of nominal shock,



pricing decisions do not depend on the anticipated duration of the price spell. While this is not the case for more general shocks, in Section 5 we show that the results and intuitions that follow do go through.

We now explore the specialized case with  $G_k(t) = \bar{G}(\Lambda_k t)$ , for some  $\bar{G}$  such that  $\int_0^\infty (1 - \bar{G}(s)) ds = 1$ . In this case, the only difference in the c.d.f. of price durations in different sectors is the scaling constant  $\Lambda_k$ . This can be interpreted as a “stretching” or “compressing” of time within each sector, with prices in sector  $k$  taking  $\frac{\Lambda_j}{\Lambda_k}$  times longer to change than in sector  $j$ .

Let  $1 - \bar{\omega}(t) = \int_t^\infty (1 - \bar{G}(s)) ds$ . It is easy to verify that:

$$\int_0^\infty (1 - G_k(s)) ds = \int_0^\infty (1 - \bar{G}(\Lambda_k s)) ds = \Lambda_k^{-1},$$

and

$$\int_t^\infty (1 - G_k(s)) ds = \int_t^\infty (1 - \bar{G}(\Lambda_k s)) ds = 1 - \bar{\omega}(\Lambda_k t).$$

To establish the role of heterogeneity in price stickiness we need a point of comparison. Following Carvalho (2006), we pick an otherwise identical one-sector economy which has the same average frequency of price changes as the multi-sector one:

**Definition 3** *The one-sector equivalent economy to a multi-sector economy characterized by the sectoral c.d.f.’s of price spells  $G_k(t) = \bar{G}(\Lambda_k t)$  with  $\int_0^\infty (1 - \bar{G}(t)) dt = 1$  and sectoral weights  $\Phi_k$  is an economy with one sector and c.d.f. of price durations given by  $\bar{G}(E[\Lambda_k]t)$ , where  $E[\Lambda_k] \equiv \sum \Phi_k \Lambda_k$ .*

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

**Proposition 3** *Let  $\Xi^{het}(t)$  denote cumulative selection of a multi-sector economy and  $\Xi^{1sect}(t)$  denote cumulative selection of its one-sector equivalent. Then,*

$$\Xi^{het}(t) < \Xi^{1sect}(t) \quad \forall t.$$

Thus, cumulative selection in the multi-sector economy is always smaller than in its one-sector equivalent. It follows immediately that a level shock to nominal aggregate demand in the multi-sector economy will have larger real effects than in its one-sector equivalent. We state this in the following corollary, which is a generalization of Carvalho (2006):

**Corollary 2** *Let  $\Gamma^{het}$  denote the real effects of a permanent level shock to nominal aggregate demand in a multi-sector economy, and  $\Gamma^{1sect}$  denote the effects of the same shock in its one-sector equivalent. Then,*

$$\Gamma^{het} > \Gamma^{1sect}.$$

The examples below help us give the intuition for the result.

**Example 1: Calvo pricing**

Consider an economy with multiple sectors associated with different average frequencies of price adjustment, and no selection within each sector. In this economy:

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

Using Lemma 3, selection in this economy is:

$$\mu^{het}(t) = \frac{1 - G^{het}(t)}{1 - \omega^{het}(t)} - 1 = \frac{\frac{E[\Lambda_k e^{-\Lambda_k t}]}{E[\Lambda_k]}}{E[e^{-\Lambda_k t}]} - 1 = \frac{cov(\Lambda_k, e^{-\Lambda_k t})}{E[\Lambda_k] E[e^{-\Lambda_k t}]} < 0. \quad (15)$$

Thus, in a heterogeneous economy with no within-sector selection, (aggregate) selection is negative. The covariance term in (15) neatly summarizes the intuition. In the heterogeneous economy, price changes are disproportionately selected from sectors with high frequency of price changes. However, exactly because these sectors have a high frequency of price changes, they have the least fraction of old prices.

In contrast, selection in the one-sector equivalent economy is zero:

$$\mu^{1sect}(t) = \frac{1 - G^{1sect}(t)}{1 - \omega^{1sect}(t)} - 1 = \frac{e^{-E[\Lambda_k]t}}{e^{-E[\Lambda_k]t}} - 1 = 0.$$

Hence, since selection is lower in the multi-sector economy than in its one-sector equivalent, the real effects of the nominal shock increase if we allow for heterogeneity in price stickiness. ■

**Example 2: Taylor pricing**

We now consider the case where the duration of price spells is constant within each sector:

$$G_k(t) = \mathbb{1}(\Lambda_k t \geq 1), \quad \omega_k(t) = \max\{\Lambda_k t, 1\},$$

so that

$$\begin{aligned} G^{het}(t) &= \frac{E[\Lambda_k \mathbb{1}(\Lambda_k t \geq 1)]}{E[\Lambda_k]} = \frac{E[\Lambda_k | \Lambda_k t \geq 1] \Pr[\Lambda_k t \geq 1]}{E[\Lambda_k]}, \\ \omega^{het}(t) &= E[\min\{\Lambda_k t, 1\}] = \Pr[\Lambda_k t \geq 1] + E[\Lambda_k t | \Lambda_k t < 1] (1 - \Pr[\Lambda_k t \geq 1]), \end{aligned}$$

with  $\Pr[\Lambda_k t \geq 1] \equiv E[\mathbb{1}(\Lambda_k t \geq 1)]$  denoting the measure of sectors where the frequency of price changes is higher than  $t^{-1}$ .

In turn, in the one-sector equivalent economy

$$G^{1sect}(t) = \mathbb{1}(E[\Lambda_k]t \geq 1), \quad \omega^{1sect}(t) = \min\{E[\Lambda_k]t, 1\}$$

It is easy to compare selection in both economies:

$$\begin{aligned} \mu^{het}(t) &= \frac{1 - \frac{E[\Lambda_k|\Lambda_k t \geq 1] \Pr[\Lambda_k t \geq 1]}{E[\Lambda_k]}}{(1 - \Pr[\Lambda_k t < 1])(1 - E[\Lambda_k|\Lambda_k t < 1]t)} - 1 \\ &= \frac{1}{1 - E[\Lambda_k|\Lambda_k t < 1]t} \frac{E[\Lambda_k|\Lambda_k t < 1]}{E[\Lambda_k]} - 1 \\ &< \frac{1}{1 - E[\Lambda_k|\Lambda_k t < 1]t} \\ &< \frac{1}{1 - E[\Lambda_k]t} \\ &= \mu^{1sect}(t). \end{aligned}$$

This shows that selection is weaker in the heterogeneous economy than in its one-sector equivalent, and thus monetary non-neutralities are larger in the former economy. ■

### 4.3 The variance of price durations

We have seen that there is a direct relationship between selection and the real effects of nominal shocks. We have also shown how this insight relates to the shape of hazard functions and the existence of cross-sectoral heterogeneity in price stickiness.

These insights are interesting for their economic content. But they are not much help for model calibration, since determining the selection or hazard functions is no easier than completely describing the distribution of the duration of price spells in a given economy.<sup>11</sup> In this section we show that, given the type of shocks and price setting environment we have been considering so far, the degree of monetary non-neutrality only depends on the average and the variance of the duration of price spells. In subsequent sections we consider more general shock processes and environments and show that the essence of the result still holds, with a few modifications.

We start by establishing a link between the variance of price spells and cumulative selection. For that purpose we use the concept of “mean preserving spread”, that is, we increase the variance of a distribution of price spells by adding to it another distribution with zero mean.

Building on results from Section 4, we start from the sequence of equalities:

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

---

<sup>11</sup>For a structural model calibrated using microeconomic estimates of the full distribution of the duration of price spells, see Vavra (2010).

If  $G_A(t)$  is obtained from a mean preserving spread of  $G_B(t)$ , then:

$$\int_0^t (1 - G_A(s)) ds \leq \int_0^t (1 - G_B(s)) ds \quad \forall t.$$

It follows immediately that:

$$\begin{aligned} e^{-\Lambda - \Lambda \Xi_A(t)} &\geq e^{-\Lambda - \Lambda \Xi_B(t)} \quad \forall t \\ \iff \Xi_A(t) &\leq \Xi_B(t) \quad \forall t. \end{aligned}$$

Note that we can reverse the steps so that if we have two economies with  $\Xi_B(t) \geq \Xi_A(t)$  and the same average duration of price spells  $\Lambda^{-1}$ , then it must be the case that  $G_A$  is a mean preserving spread over  $G_B$ .

We state the result in the following proposition:

**Proposition 4** *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 over the distribution described by  $G_B$ . 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 over  $G_B$ .*

Since economy A has lower cumulative selection at all  $t$  than economy B, it follows that the real effects of a permanent level shock to nominal aggregate demand in economy A is larger than in B.

We can in fact go further and make a one-to-one mapping between the variance of the duration of price spells and the real impact of such a monetary shock. We state the main result in the section formally in the proposition below:<sup>12</sup>

**Proposition 5** *Suppose an economy is characterized by a distribution of price spells  $G$  with finite first and second uncentered moments  $\Lambda^{-1}$  and  $\Lambda^{-2} + \sigma^2$ , where  $\Lambda$  is the frequency of price changes and  $\sigma^2$  is the variance of the duration of price spells. The real effects of a permanent level shock to nominal aggregate demand of size  $\Delta m$  satisfy*

$$\lim_{\rho \rightarrow 0} \Gamma = \frac{1}{2} (\Lambda^{-1} + \Lambda \sigma^2) \Delta m.$$

The result is striking because, for a given frequency of price changes  $\Lambda$ , the variance of price durations is a sufficient statistic for the real effects of the shock.

We now turn to particular examples which should help give intuition about the link between the variance of price durations and selection.

---

<sup>12</sup>Without linking it to selection, we first proved this result in Carvalho and Schwartzman (2008).

**Example 1: Calvo pricing**

Under Calvo pricing, the variance of the duration of price spells is  $\sigma^2 = \Lambda^{-2}$ . Thus, applying **Proposition 5**, the real effects of a shock when  $\rho \rightarrow 0$  are

$$\Gamma = \frac{\Delta m}{2} (\Lambda^{-1} + \Lambda\sigma^2) = \frac{\Delta m}{2} (\Lambda^{-1} + \Lambda^{-1}) = \Delta m \Lambda^{-1}.$$

One interesting feature of Calvo pricing is that it has zero selection, so that it implies real effects which are larger than any environment where selection is positive for a given average frequency of price change. This, in turn, implies that the variance of price durations under Calvo pricing is the largest among any environment with non-negative selection. ■

**Example 2: Taylor pricing and a perturbation**

One corollary of **Proposition 5** is that, since the constant duration model has zero variance of price durations, it also has the least real effects among all possible time-dependent pricing models for a given average frequency of price changes, confirming earlier results. Applying **Proposition 5**, the real effects of a shock under the Taylor model when  $\rho \rightarrow 0$  are

$$\Gamma = \frac{\Delta m}{2} (\Lambda^{-1} + \Lambda\sigma^2) = \frac{\Delta m}{2} (\Lambda^{-1} + 0) = \frac{\Delta m}{2} \Lambda^{-1}.$$

The dynamics of the variables of interest in response to a permanent nominal income shock are depicted in Figure 1. The lines marked  $m$  and  $p$  show the evolution of nominal income and the aggregate price level in response to a monetary shock. The difference between the two is the path of output. Nominal aggregate demand increases discretely at  $t = 0$  and the price level increases slowly. The slope of the path for the price level is given by  $\Lambda$ , the fraction of adjusting firms, multiplied by  $1 - G(t)$ , the fraction of adjusting firms which are old. Between 0 and  $\Lambda^{-1}$  this fraction is equal to one: All adjusting firms are drawn from the pool of old prices. Therefore, selection is maximal. From  $\Lambda^{-1}$  onward there are no old prices left, and thus the aggregate price level stops increasing. The real impact of the nominal shock with  $\rho \rightarrow 0$  is depicted by the shaded triangle marked  $\Gamma$  between the level of nominal aggregate demand and the price level  $p$ .

To gain more intuition about the link between the variance of price durations, selection, and the real effects of a nominal shock, we consider now the following mixture of Taylor (1979) price-setting distributions:

$$G(t) = \frac{1}{2} \mathbb{1}(t \geq \Lambda^{-1} - \Delta) + \frac{1}{2} \mathbb{1}(t \geq \Lambda^{-1} + \Delta), \quad 0 < \Delta < \Lambda^{-1}.$$

That is, a newly set price can last for an interval of length  $\Lambda^{-1} - \Delta$  or  $\Lambda^{-1} + \Delta$  with equal probability. In the mixture model the variance is higher than in the Taylor model, and is equal to  $\Delta^2 > 0$ .

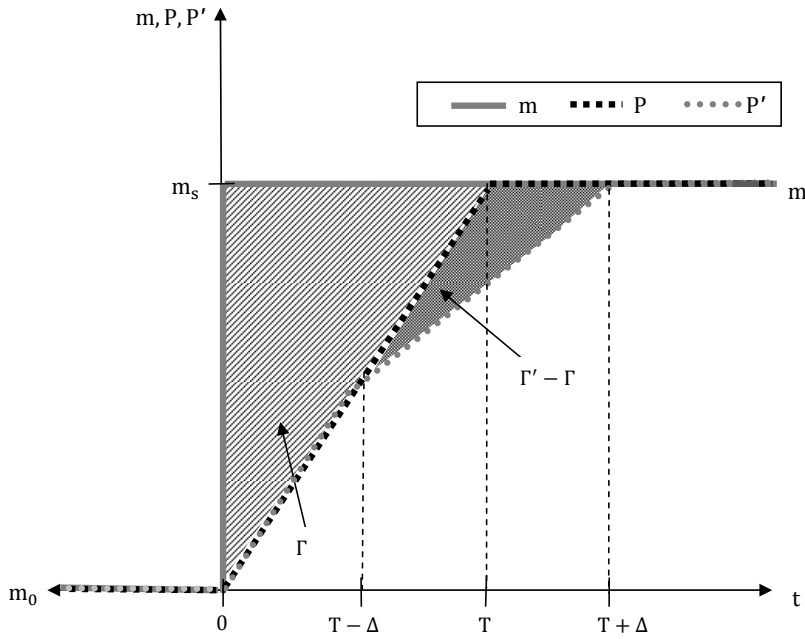


Figure 1: Response of the price level to a permanent shock to the level of nominal income.

Selection is now given by:

$$\mu(t) = \begin{cases} \frac{1}{1-\Lambda t} - 1 & \text{if } t < \Lambda^{-1} - \Delta \\ \frac{1}{1-\Lambda t + \Lambda \Delta} - 1 & \text{if } \Lambda^{-1} - \Delta < t < \Lambda^{-1} + \Delta. \end{cases}$$

Selection drops when  $t > T - \Delta$ . The intuition is straightforward. Up to  $t = \Lambda^{-1} - \Delta$  all adjusting prices come from the pool of old prices. At  $T - \Delta$  half of the prices adjusted at  $t = 0$  are up for adjustment again. The pool of adjusting prices thus also starts to include a fraction “new” prices. Selection is, therefore, less than maximal.

This example sheds light on the relationship between the variance of durations and the selection effect. An increase in the variance increases the probability that prices that were last set very recently be reset. After a shock, this means an increased probability that prices that were last set after the shock be selected again for change. This lowers selection for old prices.

Given the lower selection effect and the higher variance, the real effects of the shock under the Taylor mixture case are larger. This is evident in Figure 1. The  $p'$  line shows the response of the price level to a shock to the level of nominal aggregate demand in this mixture model. Note that the slope of  $p'$  is less steep between  $\Lambda^{-1} - \Delta$  and  $\Lambda^{-1} + \Delta$ , reflecting the weaker selection effect. In this segment, half of the firms which change prices have already responded to the shock, so that these price changes have no additional offsetting effect. For  $\rho \rightarrow 0$  we can read the difference in

the cumulative real impact of the nominal shock between the two models from the shaded triangle marked  $\Gamma' - \Gamma$ . It is immediately clear that the cumulative real effects of the shock are larger under the Taylor mixture model than under the pure Taylor. The real effects under the mixture model are:

$$\Gamma' = \frac{\Delta m}{2} (\Lambda^{-1} + \Lambda \sigma^2) = \frac{\Delta m}{2} (\Lambda^{-1} + \Lambda \Delta^2). \quad \blacksquare$$

## 5 General shocks

So far we have considered only permanent shocks to the level of nominal aggregate demand. This section provides results for more general shocks.<sup>13</sup> We retain the simplifying assumption of strategic neutrality in price setting.

One difficulty in establishing an analytical solution for general shocks is that firms which set prices after the shock will set different prices depending on when they do it. For example, if the shock leads aggregate nominal income to grow at a faster rate, firms resetting their prices later will adjust their price by a greater amount. We are able to handle these cases thanks to the following proposition, which we prove and discuss at length in a companion paper (Carvalho and Schwartzman, 2012):

**Proposition 6** *Suppose an economy is characterized by a distribution of price spells  $G$ . Suppose it starts from a deterministic path for nominal aggregate demand  $m^{old}(t)$  and at  $t = 0$  this path is subject to a one-time unexpected shock that switches it to  $m^{new}(t)$ . The real impact of the shock satisfies:*

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

The proof is in Carvalho and Schwartzman (2012). It relies on a general result that, under strategic neutrality in price setting, the real effects of a monetary shock in a generalized version of the sticky-information model of Mankiw and Reis (2002) are the same as in a sticky-price model as long as the distribution of price durations matches the distribution of the duration of price plans in the sticky-information model (and the two economies are otherwise identical in terms of all other structural features).

As in Section 4, we can write the result in terms of cumulative selection:

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

---

<sup>13</sup>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. However, in some cases we resort to shocks that do have such permanent effects as an approximation to shocks that induce persistent deviations of inflation from zero. In the context of specific models of price setting, Carvalho (2008, Appendix A.6) shows that this is appropriate as long as the discount rate ( $\rho$ ) is not strictly equal to zero.

Thus, as was the case before, the same monetary shock in two economies with the same average frequency of price changes will differ only if the distribution of price durations in these two economies are associated with different cumulative selection functions  $\Xi(t)$ . Furthermore, so long as  $m^{new}(t) > m^{old}(t)$  for all  $t$  (or vice versa) the monetary shock will have a smaller real impact in the economy with larger cumulative selection  $\Xi(t)$  everywhere. This last result implies that most of the findings from earlier sections apply to more general monetary shocks. We summarize those in the proposition below:

**Proposition 7** *Consider a shock to nominal aggregate demand characterized by  $m^{new}(t) > m^{old}(t)$  for all  $t$ .*

*I) 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:*

*I.1)  $\mu_A(t) \geq \mu_B(t) \forall t$ ,*

*I.2)  $\Xi_A(t) \geq \Xi_B(t) \forall t$ ,*

*I.3) Economy A features Taylor (1979) pricing and economy B does not.*

*I.4) Economy B features Calvo (1983) pricing and economy A features an increasing hazard function.*

*I.5) The associated hazard function  $h_A(t)$  is more increasing than  $h_B(t)$  according to the criterion in **Proposition 2**.*

*I.6)  $G_B$  is a mean-preserving spread on  $G_A$ .*

*II) Consider a heterogeneous economy with sectoral distribution of price durations  $G_k$  and sectoral weights  $\Phi_k$ . Then:*

*II.1) Let  $\Gamma^{het}$  be the cumulative real impact of the shock in a multi-sector economy. It is the same as in a one-sector economy characterized by  $G^{het}(t) = \frac{E[\Lambda_k G_k(t)]}{E[\Lambda_k(t)]}$ .*

*II.2) Suppose  $G_k(t) = \bar{G}(\Lambda_k t)$  for some  $\bar{G}$  such that  $\int_0^\infty (1 - \bar{G}(t)) dt = 1$ . Let  $\Gamma^{1sect}$  be the cumulative impact of the shock in the one sector equivalent economy defined as in **Definition 3**. Then  $\Gamma^{het} > \Gamma^{1sect}$ .*

Allowing for more general shocks uncovers additional implications of selection which are not apparent when studying level shocks. Let  $\tilde{\Delta}m \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (m^{new}(t) - m^{old}(t)) dt$  and assume that it is finite. Also, let  $\Gamma_0$  be the real effects of a level shock of size  $\tilde{\Delta}m$ . Then:

$$\lim_{\rho \rightarrow 0} \Gamma = \Gamma_0 + \int_0^\infty \left[ e^{-\Lambda t - \Xi(t)} - \frac{\Gamma_0}{\tilde{\Delta}m} \right] \left( m^{new}(t) - m^{old}(t) - \tilde{\Delta}m \right) dt,$$



where

$$\lim_{\rho \rightarrow 0} \Gamma_0 = \tilde{\Delta} m \int_0^{\infty} e^{-\Lambda t - \Xi(t)} dt.$$

$\Gamma_0$  can be interpreted as the size of the real effects in response to a level shock that matches the average shift in nominal aggregate demand. The second term summarizes the effect of having the difference between the effective and expected paths for nominal income change over time. Generally speaking, since  $e^{-\Lambda t - \Xi(t)}$  is decreasing in  $t$ , given  $\Gamma_0$ ,  $\Gamma$  is larger if  $m^{new}(t) - m^{old}(t)$  is particularly large when cumulative selection is low. This is typically soon after the shock, when most prices have not yet been reset.

We now consider the effect of permanent shocks to the growth rate of nominal aggregate demand. Such shocks are relevant for periods of disinflation such as the early 80s in the U.S. In the Appendix we present a proposition that encompasses both **Propositions 5** and **8**, by generalizing the result to any shock whose impulse response function can be well approximated by a polynomial function.

The growth rate shock is:

$$m(t) = \begin{cases} m^{old}, & t < 0 \\ m^{old} + bt, & t \geq 0. \end{cases} \quad (16)$$

**Proposition 8** *Suppose an economy is characterized by a distribution of price spells  $G$  with associated selection function  $\mu$ , hazard function  $h$  and finite first, second, and third uncentered moments  $\Lambda^{-1}$ ,  $\Lambda^{-2} + \sigma^2$ , and  $\eta\sigma^3 + 3\sigma^2\Lambda^{-2} + \Lambda^{-3}$  where  $\Lambda$  is the frequency of price changes,  $\sigma^2$  is the variance of the duration of price spells, and  $\eta$  is the skewness of the duration of price spells. The real effects of a shock to the growth rate of nominal income described by equation (16) satisfy*

$$\lim_{\rho \rightarrow 0} \Gamma = \frac{b}{6} (\Lambda^{-2} + 3\sigma^2 + \Lambda\eta\sigma^3).$$

Thus, for a given variance and mean of price durations, the real effects of a shock to the growth rate of nominal income will be larger if the distribution of price durations is skewed to the right. The intuition is that positive skewness produces a larger fraction of old prices late after the shock, when the gap between the actual and the expected paths for nominal income is largest. In terms of selection, a large skewness means that selection drops quickly as time passes. Since the shock builds over time, this low selection later in time has a disproportionate impact on the real effects of the shock.

### Example 1: Calvo pricing

Under Calvo pricing we have:

$$\begin{aligned}\sigma^2 &= \Lambda^{-2}, \\ \eta &= 2,\end{aligned}$$

so that, under a growth rate shock,

$$\begin{aligned}\lim_{\rho \rightarrow 0} \Gamma &= \frac{b}{6} (\Lambda^{-2} + 3\Lambda^{-2} + 2\Lambda\Lambda^{-3}) \\ &= b\Lambda^{-2}. \blacksquare\end{aligned}$$

### Example 2: Taylor pricing

Under Taylor pricing we have:

$$\begin{aligned}\sigma^2 &= 0, \\ \eta &= 0,\end{aligned}$$

so that, under a growth rate shock,

$$\lim_{\rho \rightarrow 0} \Gamma = \frac{b}{6} \Lambda^{-2}.$$

The real impact of a permanent growth rate shock is 6 times as large under Calvo pricing than under Taylor pricing.  $\blacksquare$

## 6 Interactions in pricing decisions

The analytical results presented so far hold under strategic neutrality in price setting. Also, so far we only considered limiting shock processes that may not be the most empirically plausible. In this section we perform some quantitative exercises to assess whether our main conclusions extend to more general cases. For the simulations, we consider the discrete time analogue of the model. In discrete time:

$$x_t = \sum (1 - G_t) [\alpha m_t + (1 - \alpha)p_t],$$

where  $1 - G_t$  is the discrete time analogue of the survival function  $1 - G(t)$ , and  $\alpha$  determines whether pricing decisions are strategic complements or strategic substitutes. In our experiments we compare the results with  $\alpha = 1$  (strategic neutrality),  $\alpha = 1/3$  (strategic complementarity) and  $\alpha = 3$  (strategic substitutability).

In order to perform the quantitative exercises, it remains to parameterize the shock process and the survival function. We follow Mankiw and Reis (2002) and consider a shock process that is

mean reverting in the growth rate of nominal aggregate demand:

$$\Delta m_t = 0.5\Delta m_{t-1} + \epsilon_t. \quad (17)$$

In order to assess the importance of higher moments in the distribution of durations of price spells, we consider first a family of survival functions which imply the same average duration of price spells but increasing variance and decreasing selection to verify the extent to which the variance of price spells affects the real effects of this more plausible shock process. Next we assess whether the third moment of price durations matters by considering a family of survival functions with the same first two moments but different third moments.

For the first experiment, the family of survival functions that we consider are truncated exponentials:

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

For each  $T$ , we pick  $\theta$  to ensure that the average duration of price spells equals 2 quarters. This family includes the two leading cases of constant duration (Taylor, 1979) and constant hazard (Calvo, 1983). The first is obtained if  $T$  coincides with the average duration. In that case,  $\theta = 0$ . The second is obtained with  $T \rightarrow \infty$ . As  $T$  increases, selection decreases and the variance of price spells increases.

Figure 2 shows the cumulative real effect of monetary shocks in (17) for different variances of price spells and for different levels of strategic interaction between firms' pricing decisions. As expected, the real effects are larger under strategic complementarity. They are also increasing in the variance of price spells, irrespective of the nature of pricing interactions. Such interactions appear to shift the lines relating the size of the real effects with the variance of price durations in an essentially parallel fashion. Selection is also quantitatively important. Moving from constant duration to constant hazard function implies an increase in real effects of similar magnitude to moving from strategic neutrality to the strategic complementarity under our chosen parameterizations.

We have verified that, irrespective of the nature of pricing interactions, the variance of the duration of price spells is an important moment to take into account when calibrating time-dependent sticky-price models. In other words, the kind of selection that we emphasize matters a great deal. We now show that incorporating information on the skewness of price durations is not important. To that end we consider a mixture of the two leading cases of constant hazard and constant duration models:

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

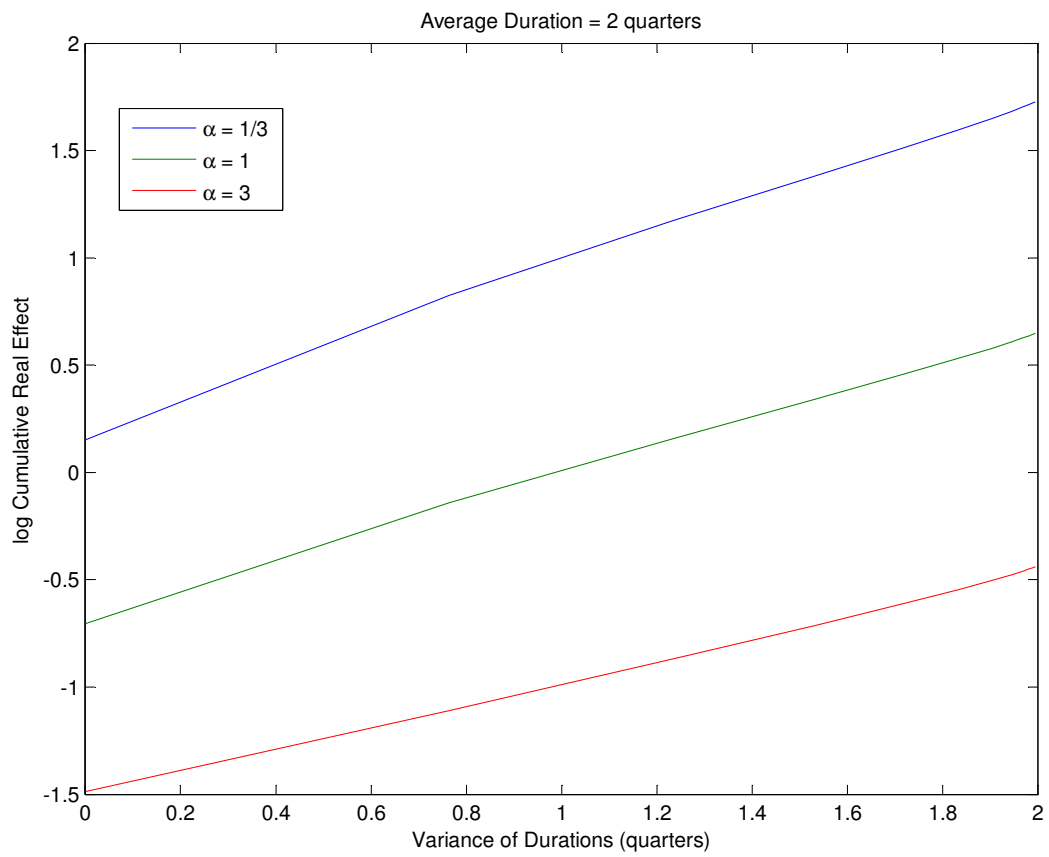


Figure 2: Response of the price level to a shock to the level of nominal income.

Under this mixture model, with probability  $q$  the survival function for a newly set price is the same as in Calvo (1983), and with probability  $1 - q$  it is the same as in Taylor (1979).

For each  $T$ , we pick  $q$  and  $\theta$  to match an average duration of price spells of 2 quarters, and a standard deviation of the duration of price spells of 1.6 quarters, corresponding roughly to the numbers reported by Klenow and Krystov (2008, page 881, Table V). Figure 3 shows, for different values of  $\alpha$ , how the real effects of the monetary shock depend on the skewness of price durations. Irrespective of the nature of pricing interactions, skewness does not seem to matter much.

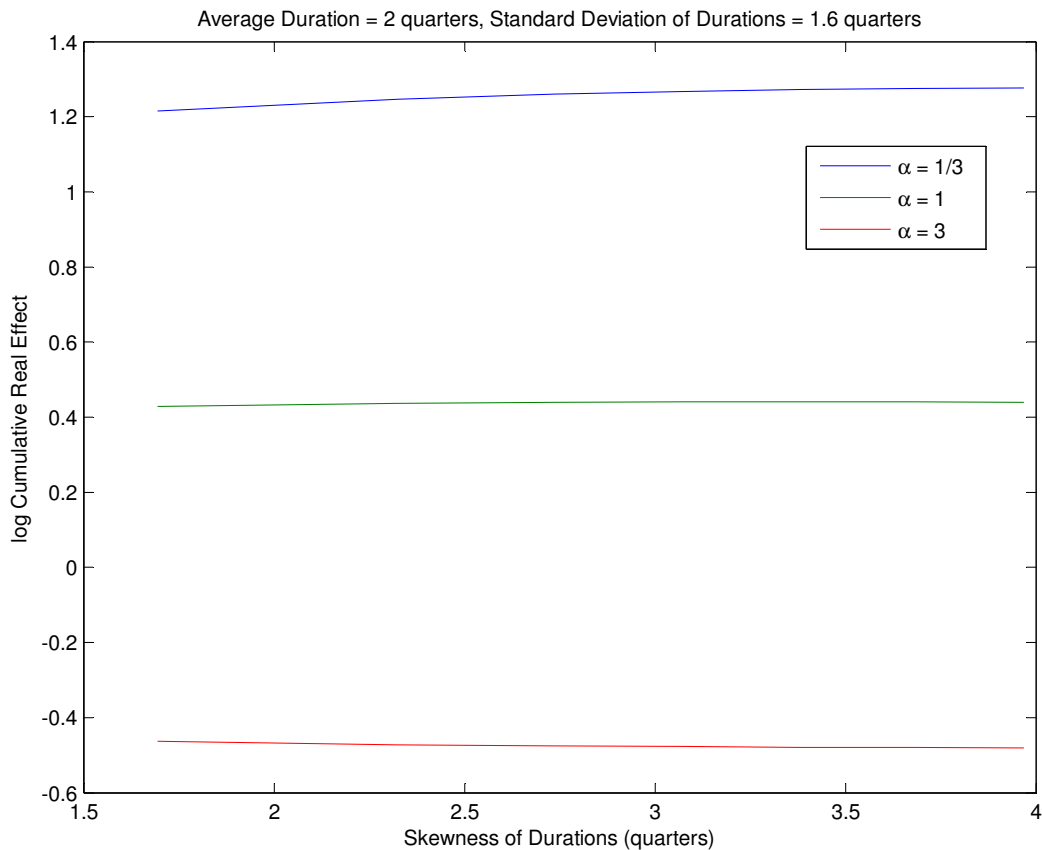


Figure 3: Response of the price level to a shock to the level of nominal income.

## 7 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 the extreme case of the Taylor (1979) model, at first all price adjustments come from firms that have not yet responded to the shock. This high degree of selection implies relatively small real effects of the nominal shock. If, however, as in Calvo (1983) or, in fact, any other time dependent scheme, there is variance in the duration of price spells, some of the price changes do not affect the price level by much, since they come from firms that have already responded to the shock.

We provide analytical expressions for the mapping between the distribution of durations of price spells and the real effects of nominal shocks under assumptions that are common in the price setting literature. Importantly, we verify that, even when they do not hold exactly, the lessons from our analysis remain relevant in more general – and arguably empirically more plausible – settings. A major take away from the paper is that when calibrating time-dependent pricing models, it is crucial to take into account not only the first, but also the second moment of the distribution of the duration of price spells. We also show that these moments should be enough under reasonable parameterizations.

While we focus on price-setting models, there is no a priori reason why many of the results and intuitions developed here could not be extended to other settings in which economic inertia stems from the fact that economic agents make decisions infrequently and in a largely time-dependent manner. These may include (but are not limited to), consumption decisions as in Lynch (1996), Gabaix and Laibson (2002) and Duffie (2010), money demand decisions as in Alvarez et al. (2009), and portfolio rebalancing problems as in Chien et al. (2010).

## References

- [1] Alvarez, F., Atkeson A. and C. Edmond (2009), “Sluggish Responses of Prices and Inflation to Monetary Shocks in an Inventory Model of Money Demand,” *Quarterly Journal of Economics* 3: 977-967.
- [2] Alvarez, F., F. Lippi, and L. Paciello (2012), “Monetary Shocks in a Model with Inattentive Producers,” mimeo.
- [3] 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.
- [4] Bils, M. and P. Klenow (2004), “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy* 112: 947-985.
- [5] Calvo, G. (1983), “Staggered Prices in a Utility Maximizing Framework,” *Journal of Monetary Economics* 12: 383-98.
- [6] Caplin A. and D. Spulber (1987), “Menu Costs and the Neutrality of Money,” *Quarterly Journal of Economics* 102: 703-726.
- [7] Carvalho, C. (2006), “Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks,” *Frontiers of Macroeconomics*: Vol. 2 : Iss. 1, Article 1.
- [8] \_\_\_\_\_ (2008), *Heterogeneity in Price Setting*, Ph.D. dissertation, Princeton University.
- [9] Carvalho, C. and F. Schwartzman (2008), “Heterogeneous Price Setting Behavior and Aggregate Dynamics: Some General Results,” mimeo available at <http://cvianac.googlepages.com/papers>.
- [10] \_\_\_\_\_ (2012), “Real Effects of Nominal Shocks in Sticky Price and Sticky Information Models: An Equivalence Result,” mimeo.
- [11] Chien, Y., Cole, H. and H. Lustig, “Is the Volatility of the Market Price of Risk due to Intermittent Portfolio Re-balancing?” mimeo available at <http://www.anderson.ucla.edu/x18962.xml>.
- [12] 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.
- [13] \_\_\_\_\_ (1999), “State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output,” *Quarterly Journal of Economics* 114: 655-690.

- [14] Duffie, D. (2010), “Asset Price Dynamics with Slow-Moving Capital” *Journal of Finance* 65: 1237-1267.
- [15] Gabaix, X. and D. Laibson (2001), “The 6D Bias and the Equity-Premium Puzzle” *NBER Macroeconomics Annual*, Vol. 16: 257-312
- [16] Kiley, M. (2002), “Partial Adjustment and Staggered Price Setting,” *Journal of Money, Banking and Credit* 34: 283-298.
- [17] 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.
- [18] Klenow, P. and B. Malin (2010), “Microeconomic Evidence on Price Setting,” *Handbook of Monetary Economics*, Ed. 1, vol. 16(6): 231 - 284.
- [19] Anthony Lynch (1996), “Decision Frequency and Synchronization Across Agents: Implications for Aggregate Consumption and Equity Return ” *Journal of Finance* 51: 1479-1497.
- [20] 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.
- [21] Sheedy (2010), “Intrinsic Inflation Persistence,” *Journal of Monetary Economics* 57: 1049-1061.
- [22] Taylor, J. (1979), “Staggered Wage Setting in a Macro Model,” *American Economic Review* 69: 108-113.
- [23] 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](https://sites.google.com/site/jvavra/docs/paper10-29-10_fordistribution.pdf?attredirects=0).
- [24] Wolman, A. (1999), “Sticky Prices, Marginal Cost, and the Behavior of Inflation,” Federal Reserve Bank of Richmond Economic Quarterly 85 (Fall 1999), 29-48.
- [25] Woodford, M. (2001), “Inflation Stabilization and Welfare,” *Contributions to Macroeconomics*: Vol. 2: Iss.1, Article 1.
- [26] \_\_\_\_\_(2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.



# 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:

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

where  $\rho$  is the discount rate and  $y_0(t)$  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, it is proportional to the total impact of the shock on the representative agent's utility. As such, it has a welfare interpretation.

To see that, log-linearize the utility function to get:

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

where  $u(t)$  is the *linear* approximation of the utility function as measured from  $t = 0$  onward,  $C$  is steady state consumption and  $L$  is steady state aggregate hours worked.

In equilibrium,  $c_j(t+s) = l_j(t+s) = y_j(t+s)$ . Also, up to a first order approximation,  $y(t) = \int y_j(t+s) dj$ . Also, in steady state,  $C = L = Y$ . Hence,

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

When comparing the utility under two different trajectories for the monetary shock we can write:

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

Thus, the utility impact of the shock is, up to a first order approximation, proportional to  $\Gamma$ . 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,

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

It follows that

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

and, since  $\varepsilon > 0$ ,  $\sigma > 0$  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  $\Gamma$  provides a useful measure of this impact.

## A.2 Proofs of propositions

**Proof of Proposition 1.** 1) and 2) follow from inspection of equation (13).

3) We check that  $f$  is given by

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

Integrating equation (11) we find that

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

From the sequence of equalities in (12) 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:

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

■

**Proof of Lemma 1.** We verify that the statement is true through a sequence of substitutions. Substituting out  $\Psi_t(s)$  we get that  $\mu(t) = \int_t^\infty \frac{h(s)}{\Lambda} \frac{1-G(s)}{\int_t^\infty (1-G(v)) dv} ds - 1$ . From equation (11) it is easy to verify that  $\Lambda \int_t^\infty (1 - G(v)) dv = 1 - \omega(t)$ . Since it does not depend on  $s$  we can take it out of the integral to get  $\mu(t) = \frac{1}{1-\omega(t)} \int_t^\infty h(s) (1 - G(s)) ds - 1$ . 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  $\mu(t) = \frac{1}{1-\omega(t)} \int_t^\infty \frac{\partial G(s)}{\partial s} ds - 1 = \frac{1-G(t)}{1-\omega(t)} - 1$ . This is exactly how  $\mu(t)$  is defined in **Definition 10**. ■

**Proof of Lemma 2.** First, note that  $\mu(0) = 0$  always, since  $G(0) = \omega(0) = 0$ . Second, we can write  $\mu(t) = E \left[ \frac{h(s)}{\Lambda} | s \geq t \right] - 1$ , where the conditional expectation is taken with respect to the probability measure  $\Psi_0$ . If the hazard function is (weakly) increasing, it follows that  $\mu(t) = E \left[ \frac{h(s)}{\Lambda} | h(s) \geq h(t) \right] - 1$ . Also, since  $h(t)$  increases in  $t$ , so does  $\mu(t)$ . Since  $\mu(0) = 0$ ,  $\mu(t) > 0 \forall t > 0$ . ■

**Proof of Proposition 2.**

1) We prove this in two steps:

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

First we show that a  $t^{**}$  with  $G_A(t^{**}) = G_B(t^{**}) < 1$  exists. Suppose not, then  $G_A(t) > G_B(t) \forall t$  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  $\Lambda_A = \int_0^\infty (1 - G_A(t)) dt$  and  $\Lambda_B = \int_0^\infty (1 - G_B(t)) dt$ .

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

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

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

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

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

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

Now, recall that  $t^{**} > t^*$ , so that if  $t > t^{**}$ , then  $h_A(t) < h_B(t)$ . 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  $1 - G_A(t) < 1 - G_B(t)$  for  $t < t^{**}$  and  $1 - G_A(t) > 1 - G_B(t)$  for  $t > t^{**}$ , then  $\Xi_A(t) > \Xi_B(t) \forall t$ .*

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

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

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:

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

Thus  $\int_0^t G_B(s) ds < \int_0^t G_A(s) ds \forall t$ . We can verify from equation (11) 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 (13),  $1 - \omega(t) = e^{-\Lambda t - \Lambda \Xi(t)}$ , 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 Corollary 1.**

Follows from applying **Proposition 1**, part 2. ■

**Proof of Lemma 3.**

See text. ■

**Proof of Proposition 3.** First, use **Definition 3** and integrate both sides of equation (11) to write  $1 - \omega^{1sect}(t)$ :

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

Now, use **Lemma 3** to write  $1 - \omega^{het}(t)$ :

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

We can show that  $\bar{\omega}$  is concave, i.e.:

$$\omega(\lambda x + (1 - \lambda)x') = \lambda \omega(x) + (1 - \lambda)\omega(x').$$

Given the definition of  $\omega$ , this is the same as:

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

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

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

Since  $\int_x^{\lambda x + (1-\lambda)x'} \frac{1}{(1-\lambda)(x'-x)} ds = \int_{\lambda x + (1-\lambda)x'}^{x'} \frac{1}{\lambda(x'-x)} ds = 1$ , the inequality holds, since

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

From Jensen's inequality, it follows that  $E[\bar{\omega}(\Lambda_k t)] < \bar{\omega}(E[\Lambda_k t])$ , and

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

The proposition follows from  $1 - \omega(t) = e^{-\Lambda - \Lambda \Xi(t)}$ . Since  $\Lambda^{het} = \Lambda^{1sect} = E[\Lambda_k]$ , and  $1 - \omega^{het}(t) > 1 - \omega^{1sect}(t)$ , then it has to be the case that  $\Xi^{het}(t) < \Xi^{1sect}(t)$ . ■

**Proof of Proposition 4.** Suppose  $G_A(t)$  is a mean preserving spread over  $G_B(t)$ . Then  $\int_0^t G_A(s) ds > \int_0^t G_B(s) ds \forall t$ . It follows that  $1 - \omega_A(t) = 1 - \Lambda \int_0^t G_A(s) ds < 1 - \Lambda \int_0^t G_B(s) ds = 1 - \omega_B(t)$ . It follows that  $\Xi_A(t) = -\frac{\ln(1-\omega_A(t))}{\Lambda} - t > -\frac{\ln(1-\omega_B(t))}{\Lambda} - t = \Xi_B(t)$ . ■

**Proof of Proposition 4.** See Carvalho and Schwartzman (2012). ■

**Proof of Proposition 7.**

I.1) and I.2) follow from parts 1) and 2) of **Proposition 1**.

I.3) Follows from the fact that, under Taylor,  $\Xi(t)$  is maximal and part I.2

I.4) Follows from **Lemma 2** and the fact that under Calvo,  $\Xi(t) = 0$  and part I.2

I.5) Follows from **Proposition 2** and part I.2.

I.6) Follows from **Proposition 4** and part I.2.

II.1) Given  $\alpha = 0$ , we can treat each sector in the multi-sector economy as a separate economy. From **Proposition 6** the real impact of the shock in an economy characterized by  $G_k$  is  $\Gamma_k = \int_0^\infty e^{-\rho t} (1 - \omega_k(t)) (m^{new}(t) - m^{old}(t)) dt$  where  $\omega_k(t) = \Lambda_k \int_0^t (1 - G_k(s)) ds$ . The real impact of the shock in the multi-sector economy is just the cross-sectoral average  $\Gamma^{het} = E[\Gamma_k] = \int_0^\infty e^{-\rho t} (1 - E[\omega_k(t)]) (m^{new}(t) - m^{old}(t)) dt$ . Now, consider 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  $\tilde{\Gamma} = \int_0^\infty e^{-\rho t} (1 - \tilde{\omega}(t)) (m^{new}(t) - m^{old}(t)) dt$  where  $\tilde{\omega}(t) = E[\Lambda_k] \int_0^t \left(1 - \frac{E[\Lambda_k G_k(s)]}{E[\Lambda_k]}\right) ds$ . It follows that  $\tilde{\Gamma} = \Gamma^{het}$  since  $E[\omega_k(t)] = E\left[\Lambda_k \int_0^t (1 - G_k(s)) ds\right] = E[\Lambda_k] \int_0^t \left(1 - \frac{E[\Lambda_k G_k(s)]}{E[\Lambda_k]}\right) ds = \tilde{\omega}(t)$ .

II.2) This follows from II.1) and **Proposition 3**. ■

**Propositions 5** and **8** are special cases of the proposition below:

**Proposition A. 4** *Suppose an economy is characterized by a distribution of prices  $G$  with  $\int_0^\infty (1 - G(t)) t^K dt < \infty$  and with associated selection function  $\mu$ , hazard function  $h$ . Let  $\tau$  be the random variable whose realizations correspond to the different price spells in the economy. The real impact of a shock to the level of nominal income characterized by  $m^{new}(t) = \sum_{k=1}^K a_k t^{k-1}$  is:*

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

**Proof.** Consider first a case with bounded support, i.e., there is  $z$  such that  $\omega(t) = 1 \forall t \geq z$ :

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

Then,

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

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

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

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  $\{z_1, z_2, \dots, z_n, \dots\}$  be an infinite sequence such that  $z_{k+1} > z_k \forall k$  and  $z_1 > 0$ . Then, for all  $t$ ,  $\lim_{n \rightarrow \infty} \Lambda_{z_n} (1 - G_{z_n}(t)) t^k = \Lambda (1 - G(t)) t^k$ . 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  $\Lambda_{z_n} = [\int_0^\infty (1 - G_{z_n}(t)) dt]^{-1}$  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  $\lim_{n \rightarrow \infty} \int_0^\infty (1 - G_{z_n}(t)) t^k dt = \int_0^\infty (1 - G(t)) t^k dt$  and  $\Lambda \int_0^\infty \sum_{k=1}^K a_k \frac{t^k}{k} (1 - G(t)) dt = \lim_{z \rightarrow \infty} \Lambda_z \int_0^\infty \sum_{k=1}^K a_k \frac{t^k}{k} (1 - G_z(t)) dt$ . ■