Intrinsic inflation persistence*

Kevin D. Sheedy†
London School of Economics

First draft: 8th June 2004
This draft: 7th October 2010

Abstract

Empirical evidence suggests that inflation determination is not purely forward looking, but models of price setting have struggled to rationalize this finding without directly assuming backward-looking pricing rules for firms. This paper shows that intrinsic inflation persistence can be explained with no deviation from optimizing, forward-looking behaviour if prices that have remained fixed for longer are more likely to be changed than those set recently. A relationship between the probability of price adjustment and the duration of a price spell is shown to imply a simple "hybrid" Phillips curve including lagged and expected inflation, which is estimated using macroeconomic data.

JEL classifications: E3.

Keywords: Inflation persistence; hazard function; time-dependent pricing; New Keynesian Phillips curve.

*I thank Petra Geraats for many helpful discussions. I am grateful to Seppo Honkapohja, Karl Whelan, Michael Woodford, Fang Yao, three anonymous referees and the editor for their comments. The paper has also benefited from the comments of seminar participants at the Bank of England, U. Birmingham, U. Cambridge, Federal Reserve Board, U. Georgia, London School of Economics, U. Manchester, New York Fed, U. Pompeu Fabra, and U. Warwick. I thank the Economic and Social Research Council for financial support, and the Journal of Applied Econometrics for an award under its Scholars Programme. An earlier draft of this paper circulated under the title "Structural Inflation Persistence".

†Department of Economics, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK. Tel: +44 207 107 5022, Fax: +44 207 955 6592, Email: k.d.sheedy@lse.ac.uk, Website: http://personal.lse.ac.uk/sheedy.

1 Introduction

Empirical evidence suggests that inflation determination is not purely forward-looking, in that recent past inflation rates have a role in explaining inflation outcomes in addition to that of current and expected future economic fundamentals.[1] This is reflected in many current DSGE models using Phillips curve equations that include past inflation rates (Smets and Wouters, 2003, Christiano, Eichenbaum and Evans, 2005). In spite of the wide application of such Phillips curves, providing a coherent theoretical rationale for intrinsic inflation persistence has proved challenging.[2]

The widely used Calvo (1983) pricing model assumes that firms are randomly selected to make price changes, and this leads to a theory of inflation (the New Keynesian Phillips curve) in which history is irrelevant once current and expected future fundamentals (such as real marginal cost or the output gap) are known. It has been necessary to bolt on extra ad hoc assumptions to this theory in order to obtain Phillips curves including past inflation. Smets and Wouters (2003) and Christiano, Eichenbaum and Evans (2005) make use of a "dynamic indexation" assumption whereby all firms' prices are continually adjusted according to a rule making use of past inflation data. Galí and Gertler (1999) suggest that some fraction of firms uses a "rule-of-thumb" to raise prices mechanically in line with past inflation rates.

This paper argues that the role of past inflation can be rationalized instead simply by discarding the assumption of random selection of those firms that change price. If prices are more likely to be reviewed and changed after a prolonged spell of stickiness then this can immediately explain the presence of lagged inflation in the Phillips curve with a positive coefficient. This alternative is arguably less ad hoc since there is no reason to believe that all firms have equal incentives to change price irrespective of the time since the previous adjustment, and good reasons to believe those that have waited longer will have more to gain from adjustment, other things being equal. This "selection effect" based on the duration of price stickiness thus provides a simple theory of how optimizing, forward-looking behaviour by firms can explain the intrinsic persistence of inflation.[3]

The intuition for the consequences of the selection effect is straightforward. Consider a cost-push shock, or an increase in demand pressure as measured by the output gap, lasting for just one time period. With staggered price adjustment, some prices remain unchanged, while others rise, increasing the aggregate price level. After the shock has dissipated there are two groups of firms and two countervailing effects on aggregate inflation. The firms that did not initially respond now find that although the original shock has gone, the aggregate price level has increased, so they want to raise their prices in money terms. This is the "catch-up" effect. The firms that did respond must necessarily have made price increases on average that exceeded the aggregate

[1] There is an extensive literature on this question, starting from Fuhrer and Moore (1995). Important contributions include Galí and Gertler (1999), Roberts (2005) and Rudd and Whelan (2005). The evidence on inflation persistence in general, and intrinsic persistence in particular, is surveyed by Fuhrer (2010).

[2] It has been argued that deviations from rational expectations (Roberts, 1997, Paloviita, 2004) or adaptive learning (Milani, 2005) could account for intrinsic inflation persistence. Furthermore, it might be the case that time-variation in the average rate of inflation generates apparent inflation persistence (Cogley and Sbordone, 2005). There is further discussion of these ideas in Woodford (2007).

[3] If the selection effect works in the opposite direction, in that prices set more recently are more likely to be reviewed and changed again, then past inflation rates still appear in the Phillips curve, but now with negative coefficients.

inflation rate. This means that once the shock has gone, their money prices are now too high and they would like to reduce them. This is the "roll-back" effect. A selection effect whereby firms whose prices have remained unchanged for longer are now more likely to change them increases the importance of "catch-up" relative to "roll-back" when compared to the case of purely random selection. When "catch-up" dominates "roll-back", aggregate inflation remains positive even though the original shock has gone. This gives rise to inflation persistence that cannot be explained in terms of persistence in the fundamentals driving inflation, in other words, intrinsic inflation persistence.[4]

This paper makes three distinct contributions to the literature on inflation dynamics and price setting. First, the paper makes a theoretical contribution in showing how the slope of the hazard function (the relationship between the probability of price adjustment and the existing length of a price spell) determines whether intrinsic inflation persistence is present. An upward-sloping hazard function generates such intrinsic persistence; a downward-sloping hazard function does not (and actually implies inflation displays less persistence than that present in the underlying fundamentals). These findings can be understood in terms of a selection effect not present in a model with a flat hazard function.

Second, the paper makes a methodological contribution. It derives simple "hybrid" Phillips curves (that is, Phillips curves containing both lagged and expected future inflation) in cases where the probability of price adjustment does depend on the existing length of a price spell. The representation of the Phillips curve obtained here for non-constant hazard functions is easier to analyse than that used in the existing literature, to which it is observationally equivalent. The widely used New Keynesian Phillips curve is a special case of this representation, though its advantage of simplicity prized in many applications is shown here to be available in a much broader class of models. This methodological innovation may have applications in applied DSGE analyses where hybrid Phillips curves (based on backward-looking pricing rules) have proved popular. It also has applications in optimal monetary policy analyses that relax the assumption of Calvo pricing (as has been done by Khan, King and Wolman, 2003). The simple Phillips curve representation is used by Sheedy (2007) to derive the implications for optimal monetary policy of the selection effect and the resulting intrinsic inflation persistence (see also Woodford, 2010, for an exposition of the results).

Third, the paper makes an empirical contribution in estimating the hazard function consistent with U.S. inflation dynamics. This is done by developing a method for identifying and estimating the hazard function using only macroeconomic data and simple single-equation econometric techniques — no individual price observations are required.[5] The familiar method of obtaining an overall frequency of price adjustment from estimation of the New Keynesian Phillips curve is a special case

[4]This intuition generalizes to the case of persistent shocks: with a positive shock, prices set further in the past are on average lower than those set more recently, and thus will rise by more when they are changed. Therefore any increase in the likelihood of changing older prices relative to newer prices leads to a higher rate of inflation in the periods following the shock than would otherwise occur.

[5]Estimates of hazard functions using macroeconomic data are rare owing to the econometric difficulties these entail. Jadresic (1999) provides estimates, but using an OLS method that is valid only when all expectations of inflation or the output gap conditioned on different information sets are identical. This requires a strong perfect foresight assumption. The techniques developed here require only the conventional rational expectations assumption. Laforte (2007) estimates a version of the Wolman (1999) model, featuring a non-constant hazard function, as part of a DSGE model estimated using Bayesian methods.

of the procedure developed here. Using this approach, it is possible to test whether there exists a well-defined hazard function consistent with observed inflation dynamics when firms set prices in a purely forward-looking manner. Hazard functions are found to be largely upward sloping, albeit with a dip and subsequent rise around a price-spell duration of one year. These estimated hazard functions can be compared to those from the burgeoning microeconometric literature.[6]

This paper is of course not the first to consider a model of price adjustment with a non-constant hazard function. The original Taylor (1980) contracting model features fixed-length price spells, which is a special case of an increasing hazard. The implications of Taylor contracts for inflation dynamics are analysed by Guerrieri (2001, 2002). A general hazard function is considered by Goodfriend and King (1997), while Wolman (1999) argues for an increasing hazard function on the grounds that this is an implication of models of state-dependent pricing (for example, Dotsey, King and Wolman, 1999). A mixed Calvo-Taylor specification is considered by Mash (2004). Laforte (2007) performs a Bayesian estimation of a DSGE model with a range of competing pricing assumptions: Calvo pricing with indexation, sticky information (Mankiw and Reis, 2002), and the increasing-hazard specification of Wolman (1999), finding that the data favour the Wolman model over the alternatives. These studies have concluded that upward-sloping hazard functions are better able to match the empirical impulse response and autocorrelation functions of inflation. However, the logic behind these findings has not been fully understood because none of these papers analyses the consequences of downward-sloping hazards, where the opposite conclusion is found. The success of the alternative models cannot be explained simply in terms of deviating from the overly restrictive Calvo pricing assumption: it is necessary to move specifically in the direction of an upward-sloping hazard. Moving toward a negatively sloped hazard function actually delivers an even worse performance in matching inflation dynamics.

In spite of the better empirical performance of models with upward-sloping hazard functions, a number of papers have questioned whether these are really consistent with intrinsic inflation persistence. Whelan (2007) and Yao (2009) have derived Phillips curves purporting to show that all hazard functions, whether upward- or downward-sloping, are inconsistent with lagged inflation rates having positive coefficients (echoing the earlier findings of Fuhrer (1997) in the case of Taylor contracts). The Phillips curves derived in these papers include moving averages of fundamentals (real marginal cost or the output gap) and past expectations of inflation and the fundamentals (as in "sticky information" models such as Mankiw and Reis, 2002), in addition to lags of inflation themselves. These Phillips curves thus differ fundamentally from the usual "hybrid" Phillips curves considered in the empirical literature. This paper shows that these extra terms are observationally equivalent to lagged inflation having a positive coefficient if and only if the hazard function is upward sloping. Such a finding is consistent with Dotsey (2002), who shows that data generated

[6]The extensive microeconometric literature is surveyed in Klenow and Malin (2010). Micro-data estimates of the hazard function tend to be largely flat or downward sloping, with the exception of a spike at the one-year duration (Nakamura and Steinsson, 2008). Some studies do find evidence of upward-sloping hazard functions, for example, Cecchetti (1986), Götte, Minsch and Tyran (2005), Ikeda and Nishioka (2007), and Cavallo (2009). Álvarez, Burriel and Hernando (2005) argue that the common finding of a downward-sloping hazard function may be due to a heterogeneity bias in the microeconometric estimates. Eichenbaum, Jaimovich and Rebelo (2008) find mildly upward-sloping hazards by excluding non-"reference prices".

by a model with Taylor contracts would lead to an estimated hybrid Phillips curve with positive intrinsic inflation persistence. The finding is also in line with Bakhshi, Khan and Rudolf (2007), who find a positive coefficient on lagged inflation in the context of a model with an upward-sloping hazard function.

The plan of the paper is as follows. The model is set out in section 2 and the simple representation of the Phillips curve is derived there. A parameterization of the hazard function capturing the selection effect is developed in section 3. Section 4 shows how the selection effect determines the sign of the coefficients of past and expected future inflation in the Phillips curve. Section 5 presents the macro-data estimation strategy and the resulting hazard functions. Finally, section 6 draws some conclusions.

2 Price setting and the Phillips curve

The economy contains a measure-one continuum Ω of firms producing differentiated goods. Preferences over goods are given by a CES aggregator with elasticity of substitution ε ($\varepsilon > 1$). The demand function faced by firm $\imath \in \Omega$ is

$$\mathsf{y}_t(\imath) = -\varepsilon(\mathsf{p}_t(\imath) - \mathsf{p}_t) + \mathsf{y}_t, \qquad [2.1]$$

where $\mathsf{y}_t(\imath)$ is log output and $\mathsf{p}_t(\imath)$ the log price of the good produced by firm $\imath$. Log aggregate output is $\mathsf{y}_t = \int_\Omega \mathsf{y}_t(\imath)d\imath$, and the log general price level is $\mathsf{p}_t = \int_\Omega \mathsf{p}_t(\imath)d\imath$. The price elasticity of demand is ε. All variables (except prices) are given as deviations from their steady-state values, and all equations are given as log linearizations of the original non-linear equations where appropriate.[7]

The log real marginal cost of production for firm $\imath$ is

$$\mathsf{x}_t(\imath) = \eta_\mathrm{f}\mathsf{y}_t(\imath) + \eta_\mathrm{y}\mathsf{y}_t + \mathsf{z}_t, \qquad [2.2]$$

where η_f is the elasticity of marginal cost with respect to the firm's own output ($\eta_\mathrm{f} \geq 0$), η_y is the elasticity with respect to aggregate output ($\eta_\mathrm{y} > 0$), and z_t denotes any aggregate factors other than output that affect costs.[8] With the constant price-elasticity demand function implied by the CES aggregator, the profit-maximizing (log) price $\mathsf{p}^*_t(\imath)$ for a firm with fully flexible prices is to set a constant markup on its marginal cost, hence $\mathsf{p}^*_t(\imath) = \mathsf{p}_t + \mathsf{x}_t(\imath)$. By combining this with equations [2.1] and [2.2], it follows that the profit-maximizing flexible price is the same for all firms and is given by

$$\mathsf{p}^*_t = \mathsf{p}_t + \nu\mathsf{x}_t, \quad \text{where } \nu \equiv \frac{1}{1+\varepsilon\eta_\mathrm{f}}, \qquad [2.3]$$

and $\mathsf{x}_t = \int_\Omega \mathsf{x}_t(\imath)d\imath$ is real marginal cost averaged across all firms, which is $\mathsf{x}_t = (\eta_\mathrm{f} + \eta_\mathrm{y})\mathsf{y}_t + \mathsf{z}_t$.

[7]These log linearizations are standard. For further details, see Wolman (1999) or Woodford (2003). The steady state chosen here features zero real GDP growth and zero inflation for simplicity, though relaxing either of these assumptions to allow for moderate growth or inflation would not significantly alter the results.

[8]For the purposes of this paper it is more convenient directly to specify the marginal cost function rather than the production function from which it is derived.

The coefficient ν ($\nu > 0$) is inversely related to the extent of real rigidity in the sense of Ball and Romer (1990), and is also negatively related to strategic complementarity in pricing decisions in some contexts.

Price setting is modelled using the framework of time-dependent price adjustment. The key ingredient is the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$, where α_ℓ is the probability of a firm now changing a price set ℓ periods ago. The hazard function implies a survival function $\{\psi_\ell\}_{\ell=0}^{\infty}$, where ψ_ℓ is the probability a newly set price will still be in use in ℓ periods' time. The survival function is calculated from the hazard function using $\psi_\ell = \prod_{i=1}^{\ell}(1-\alpha_i)$. A price is used for at least one period, thus $\psi_0 = 1$.

This paper focuses on hazard functions implying some minimal degree of price stickiness, that is, $\alpha_1 < 1$, but where all prices will eventually adjust, which is ensured by $\alpha_\infty > 0$ (with $\alpha_\infty \equiv \lim_{\ell\to\infty} \alpha_\ell$). If a firm chooses a new price at time t (its reset price) and expects future probabilities of changing price are determined by the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ then the profit-maximizing reset price r_t is a weighted average of current and expected future p_t^*'s:

$$\mathsf{r}_t = \sum_{\ell=0}^{\infty} \left(\frac{\beta^\ell \psi_\ell}{\sum_{i=0}^{\infty} \beta^i \psi_i} \right) \mathbb{E}_t \mathsf{p}_{t+\ell}^*, \qquad [2.4]$$

where β is the discount factor (satisfying $0 < \beta < 1$). All firms changing price at the same time choose the same reset price.

The general price level p_t is an aggregate of current and past reset prices. If $\omega_{\ell,t}$ denotes the proportion of firms at time t using a price set ℓ periods ago then $\mathsf{p}_t = \sum_{\ell=0}^{\infty} \omega_{\ell,t}\mathsf{r}_{t-\ell}$. The evolution of the age distribution of prices $\{\omega_{\ell,t}\}_{\ell=0}^{\infty}$ is calculated using the hazard function with equations $\omega_{\ell,t} = (1-\alpha_\ell)\omega_{\ell-1,t-1}$ and $\omega_{0,t} = \sum_{\ell=1}^{\infty} \alpha_\ell \omega_{\ell-1,t-1}$.

Proposition 1 *There is a unique stationary age distribution of prices, that is, a distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$ such that $\omega_{\ell,t} = \omega_\ell$ for all t and all ℓ. If the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ is bounded away from zero, that is, $\alpha_\ell \geq \underline{\alpha}$ for all ℓ and some $\underline{\alpha} > 0$, then there is convergence to the stationary age distribution ($\omega_{\ell,t} \to \omega_\ell$ as $t \to \infty$) starting from any initial age distribution $\{\omega_{\ell,t_0}\}_{\ell=0}^{\infty}$ at time t_0.*

Proof See appendix A.2. ■

The stationary age distribution has the property that $\omega_\ell = (1-\alpha_\ell)\omega_{\ell-1}$, so $\omega_\ell = \psi_\ell\omega_0$. If the economy is at the stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$ then the general price level p_t is:

$$\mathsf{p}_t = \sum_{\ell=0}^{\infty} \omega_\ell \mathsf{r}_{t-\ell}. \qquad [2.5]$$

In what follows, it is assumed that the economy is always at the stationary age distribution.

It is possible to derive a Phillips curve relationship between inflation $\pi_t \equiv \mathsf{p}_t - \mathsf{p}_{t-1}$ and real marginal cost x_t (and hence aggregate output y_t, or the output gap) by directly combining equations

[2.3], [2.4] and [2.5]:

$$\pi_t = \sum_{\ell=1}^{\infty} a_\ell \pi_{t-\ell} + \sum_{i=0}^{\infty}\sum_{\ell=1}^{\infty} b_{i\ell}\mathbb{E}_{t-i}\pi_{t-i+\ell} + \nu \sum_{i=0}^{\infty}\sum_{\ell=0}^{\infty} c_{i\ell}\mathbb{E}_{t-i}x_{t-i+\ell}, \qquad [2.6]$$

where the coefficients a_ℓ, $b_{i\ell}$ and $c_{i\ell}$ depend on the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ and the discount factor β. A Phillips curve of this form has been derived and studied in a number of papers (Mash, 2004, Whelan, 2007, Yao, 2009). The Phillips curve [2.6] contains terms in expected future inflation, past inflation, and past expectations of current, future and past inflation, along with a similar set of terms in past and expected future real marginal cost, and with a full set of past expectations of real marginal cost as well. This representation of the Phillips curve is unnecessarily complicated for both theoretical and empirical analysis, so an observationally equivalent simpler representation is presented below.

Proposition 2 *The Phillips curve [2.6] with any hazard functions for which there is convergence to the stationary age distribution is observationally equivalent to a Phillips curve of the alternative form:*

$$\pi_t = \sum_{\ell=1}^{\infty} \lambda_i \pi_{t-\ell} + \sum_{\ell=1}^{\infty} \xi_i \mathbb{E}_t \pi_{t+\ell} + \nu\kappa x_t, \qquad [2.7]$$

for some coefficient κ and sequences of coefficients $\{\lambda_\ell\}_{\ell=1}^{\infty}$ and $\{\xi_\ell\}_{\ell=1}^{\infty}$ with $\sum_{\ell=1} |\lambda_\ell| < \infty$ and $\sum_{\ell=1}^{\infty} |\xi_\ell| < \infty$.

Proof See appendix A.3. ■

The alternative representation of the Phillips curve exists if and only if the hazard function implies that there will always be convergence to the stationary age distribution. This rules out hazard functions that are consistent with stable cycles in the age distribution, for example, when firms synchronize on price adjustment in certain periods. A weak sufficient condition for the existence of the alternative representation [2.7] is that the hazard function is never exactly equal to zero (as can be seen from Proposition 1 in combination with Proposition 2). Microeconometric estimates of hazard functions suggest this condition should hold in practice[9], though it does rule out certain theoretical special cases, including Taylor (1980) contracts. The difficulty here with Taylor contracts is *not* that they imply a spike in the hazard function, *nor* that the hazard function rises to probability one exactly, but that the hazard is exactly at zero until the end of the contract is reached. This zero probability means that the hazard function is consistent with both staggering and synchronization of price changes, hence there is no guarantee of convergence to the stationary age distribution. However, there are always hazard functions that do imply convergence arbitrarily close to any hazard function that does not, so the conditions required for the existence of Phillips curve representation [2.7] are not very onerous.[10]

[9] See Nakamura and Steinsson (2008), for example.

[10] For example, a hazard function specifying a fixed number of periods with a very small probability of adjustment per period, and then an adjustment probability equal to one afterwards, can be made arbitrarily close to a Taylor contract, but which always implies convergence to the stationary age distribution.

It should be noted the reason why convergence to the stationary distribution is required in Proposition 2 is *not* to ensure the coefficients in the price-level equation [2.5] are time invariant. It is always possible to contemplate starting from the stationary distribution even though there is no guarantee of convergence, and this is exactly what was assumed in obtaining the standard Phillips curve [2.6]. The substantive content of Proposition 2 is that when there is convergence to the stationary age distribution, equations [2.4] and [2.5], which involve moving averages of future and past variables, can be replaced by equivalent "autoregressive" (or recursive) equations. This then allows the simpler Phillips curve representation [2.7] to be derived. The full proof of this claim is in appendix A.3, with the key steps in the argument outlined below.

Sketch proof To obtain the Phillips curve [2.7] it will be necessary to replace the price-level equation $\mathsf{p}_t = \sum_{\ell=0}^{\infty} \omega_\ell \mathsf{r}_{t-\ell}$ with an equation of the form

$$\mathsf{p}_t = \sum_{\ell=1}^{\infty} \phi_\ell \mathsf{p}_{t-\ell} + \left(1 - \sum_{\ell=1}^{\infty} \phi_\ell\right) \mathsf{r}_t, \qquad [2.8]$$

for some sequence of coefficients $\{\phi_\ell\}_{\ell=1}^{\infty}$ (the expression for the coefficient of p_t above is required for consistency with $\sum_{\ell=0}^{\infty} \omega_\ell = 1$). An "autoregressive" equation of this form implying the same relationship between r_t and p_t as the original equation will only be well defined if the sequence of coefficients $\{\phi_\ell\}_{\ell=1}^{\infty}$ is (absolutely) summable. The condition required for this is that the "moving-average" equation [2.5] be *invertible*, that is, all the roots of the lag polynomial for the moving average lie strictly outside the unit circle.[11] Hence, any root $z = \zeta^{-1}$ of $\sum_{\ell=0}^{\infty} \omega_\ell z^\ell = 0$ must satisfy $|\zeta| < 1$.

It turns out this condition is essentially the same as that for there to be convergence to the stationary age distribution. Note that since $\omega_{0,t} = 1 - \sum_{\ell=1}^{\infty} \omega_{\ell,t}$, it follows that the transition equations for the age distribution can be written as

$$\omega_{1,t} = (1-\alpha_1) - (1-\alpha_1)\sum_{\ell=1}^{\infty} \omega_{\ell,t-1}, \quad \text{and} \quad \omega_{\ell,t} = (1-\alpha_\ell)\omega_{\ell-1,t-1} \;\; \text{for} \;\; \ell = 2, 3, \ldots.$$

Now consider any eigenvalue ζ of the linear mapping from the sequence $\{\omega_{\ell,t-1}\}_{\ell=1}^{\infty}$ to $\{\omega_{\ell,t}\}_{\ell=1}^{\infty}$. This eigenvalue satisfies $\zeta \mathrm{v}_1 = -(1-\alpha_1)\sum_{\ell=1}^{\infty} \mathrm{v}_\ell$ and $\zeta \mathrm{v}_\ell = (1-\alpha_\ell)\mathrm{v}_{\ell-1}$ for $\ell = 2, 3, \ldots$, where the sequence $\{\mathrm{v}_\ell\}_{\ell=1}^{\infty}$ is a corresponding eigenvector. If $\zeta \neq 0$ then these equations imply

$$\mathrm{v}_1 + \frac{(1-\alpha_1)}{\zeta}\sum_{\ell=1}^{\infty} \mathrm{v}_\ell = 0, \quad \text{and} \;\; \mathrm{v}_\ell = \frac{(1-\alpha_\ell)}{\zeta}\mathrm{v}_{\ell-1} \;\; \text{for} \;\; \ell = 2, 3, \ldots.$$

It follows that $\mathrm{v}_\ell = \mathrm{v}_1 \prod_{i=1}^{\ell}((1-\alpha_i)/\zeta)$, which allows $\mathrm{v}_2, \mathrm{v}_3, \ldots$ to be eliminated from the equation above, yielding:

$$\left\{ \sum_{\ell=0}^{\infty} \frac{\prod_{i=1}^{\ell}(1-\alpha_i)}{\zeta^\ell} \right\} \mathrm{v}_1 = 0.$$

[11] For a discussion of the concept of invertibility in the context of ARMA models, see Hamilton (1994).

If it were the case that $v_1 = 0$ then v_ℓ is necessarily equal to zero for all ℓ, which would mean that $\{v_\ell\}_{\ell=1}^{\infty}$ cannot be an eigenvector. This allows $v_1 \neq 0$ to be cancelled from the equation above. Now note that the stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$ satisfies $\omega_\ell/\omega_0 = \psi_\ell$ with $\omega_0 \neq 0$, and that $\psi_\ell = \prod_{i=1}^{\ell}(1-\alpha_i)$. Therefore, any eigenvalue ζ must satisfy $\sum_{\ell=0}^{\infty}\omega_\ell(\zeta^{-1})^\ell = 0$.

This argument demonstrates there is an equivalence between (non-zero) eigenvalues of the transition equations for the age distribution and reciprocals of (finite) roots of the lag polynomial for equation [2.5]. Since invertibility of the lag polynomial requires reciprocals of all these roots to have modulus less than one, this is seen to be equivalent to requiring all eigenvalues to have modulus less than one, which is in turn equivalent to there being convergence to the stationary age distribution.■

The observationally equivalent Phillips curve [2.7] contains only current real marginal cost, expected future inflation, and past inflation. Note that the representation [2.7] is contained within the class [2.6], so it is simpler in general. The fact that all expectations in [2.7] are taken with respect to the same time-t information set, together with the absence of lags and leads of real marginal cost, makes equation [2.7] an easier representation of the Phillips curve to analyse and estimate.

The Calvo (1983) pricing model, which assumes the same constant probability of price adjustment for all firms ($\alpha_\ell = \alpha$), is well known to lead to the New Keynesian Phillips curve (NKPC) $\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \nu\kappa x_t$ (see Woodford, 2003, for further details). The NKPC is a special case of [2.7] with $\xi_1 = \beta$ and all other ξ_ℓ and λ_ℓ coefficients being zero. This model of the Phillips curve is widely used owing to its tractability. Calvo pricing also implies a Phillips curve of the form [2.6], but the complexity of this equation means that it is never used in practice. This paper establishes that the simpler form [2.7] of the Phillips curve is available for a very wide class of time-dependent pricing models, namely all those where there is convergence to the stationary age distribution (a weak sufficient condition for which is given in Proposition 1).

Equation [2.7] has a close resemblance to the hybrid Phillips curves obtained by adding ad hoc persistence-generating mechanisms to the Calvo model, for example, "rule-of-thumb" firms (Galí and Gertler, 1999), or "dynamic indexation" (Christiano, Eichenbaum and Evans, 2005, Smets and Wouters, 2003). It is important to note that the lags of inflation in [2.7] arise not because of such features, but only as a result of deviations from the constant hazard rate imposed by Calvo pricing. Non-Calvo models of time-dependent pricing have been studied before, but by using Phillips curves of the more complicated form [2.6].

The simpler form [2.7] of the Phillips curve has important advantages in understanding the link between inflation dynamics and features of the hazard function. This is because part of its simplicity stems from it being a purely *autoregressive* equation in inflation, while the standard Phillips curve [2.6] also includes moving averages of real marginal cost. Just as both autoregressive and moving-average components of an ARMA stochastic process contribute to the implied patterns of serial correlation, the dynamics of inflation implied by [2.6] depend on the leads and lags of real marginal cost as well as the leads and lags of inflation.

In equation [2.6], the mapping from the coefficients to the implied inflation dynamics is made even more complicated by the presence of lags of expectations, which also make a distinct contribution to inflation dynamics. For example, in the "sticky information" model of Mankiw and Reis (2002), the dynamics of inflation are determined solely by these lagged expectations of inflation and real marginal cost.

The new representation [2.7] of the Phillips curve avoids these problems of interpretation by having no "moving average" component, and no lagged expectations. The argument of Proposition 2 is that convergence to the stationary age distribution is equivalent to the "moving average" components of [2.6] being invertible in the sense that there is a purely autoregressive observationally equivalent equation. Furthermore, the proposition shows that working with the purely autoregressive form avoids having lagged expectations in the equation.[12]

The ability to deduce inflation dynamics easily from [2.7] has important advantages. Existing theoretical studies have focused on the properties of the Phillips curve representation [2.6], but the findings are hard to interpret for the reasons discussed above. For example, it is known that all the coefficients of lagged inflation a_ℓ in [2.6] are negative, irrespective of the shape of the hazard function. However, as will be seen below, this finding has no implications whatsoever for the signs of lagged inflation in the observationally equivalent Phillips curve [2.7]. Furthermore, the empirical literature on hybrid Phillips curves has estimated equations of the form [2.7], not [2.6], and estimated coefficients in one equation reveal little in general about the coefficients of the corresponding terms in the alternative equation.

3 Hazard functions and the duration selection effect

Having established that the Phillips curve can be represented as an equation of the form [2.7] for a general hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ implying convergence to the stationary age distribution, the next step is to understand how the coefficients on lagged and future inflation are affected by features of the hazard function. The case of Calvo pricing is useful as a benchmark here because its distinctive assumption is of entirely random selection of the firms that change price. Calvo pricing also entails a constant fraction of firms changing price at any time, but this feature is shared by all time-dependent pricing models (at the stationary age distribution), and even by some state-dependent pricing models (Danziger, 1999, Gertler and Leahy, 2008).

As argued by Golosov and Lucas (2007), the correspondence between price stickiness at the microeconomic level and rigidity of the aggregate price level can be radically changed by the presence of a "selection effect", whereby those firms whose prices are far from what would maximize profits are the ones that change price. Here, this paper studies a distinct, but related, selection effect: whether the firms that change price are disproportionately drawn from among those that have not made a price change for a long time, or from those that have changed very recently.[13] In other words, this selection effect works through the duration of a price spell and is thus linked to the shape of the hazard function. A positive selection effect (positively sloped hazard function) means

[12]Since the autoregressive equations have a recursive form, the order of the lag operator and the expectations operator will be reversed when compared to the derivation of the Phillips curve using the "moving-average" equations.

[13]In a model of (homogeneous) time-dependent pricing, this is the only possible type of selection effect.

that the probability of changing price in a given period is higher for older prices than newer ones, while a negative selection effect (negatively sloped hazard function) means that more recently set prices are more likely to be changed again in a given time period. The absence of a selection effect is equivalent to an entirely flat hazard function.

There is a natural reason for supposing the duration selection effect is positive. The average gains from changing price are likely to be larger the longer a price has been left unadjusted. For example, non-stationarity in the price (in money terms) that a firm considers desirable would give rise to a positive selection effect (the argument works for both deterministic and stochastic trends).[14]

In contrast to the selection effect of Golosov and Lucas (2007), here the emphasis is not primarily on the rigidity of the general price level, but instead on how the duration selection effect determines the extent of intrinsic inflation persistence: the degree to which inflation depends on its own past, independently of current and expected future fundamentals.

The consequences of the duration selection effect are studied by parameterizing the hazard function so that there are parameters specifically controlling the direction and strength of the selection effect independently of the degree to which individual prices are flexible on average. A parameter α will determine the overall frequency of price adjustment, while n parameters $\{\varphi_i\}_{i=1}^{n}$ will determine the direction and strength of the selection effect ($n \geq 1$). Equivalently, α will determine the average level of the hazard function, while $\{\varphi_i\}_{i=1}^{n}$ will determine its slope. The number n of selection-effect parameters may be increased to expand the range of possible hazard-function shapes that can be accommodated, though attention will focus on parameterizations with a finite number of parameters.

It will be seen that the parameters α and $\{\varphi_i\}_{i=1}^{n}$ have the interpretations given above when the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ is defined recursively by

$$\alpha_\ell = \alpha - \sum_{i=1}^{n} \varphi_i + \sum_{i=1}^{\min\{\ell-1,n\}} \varphi_i \left(\prod_{j=\ell-i}^{\ell-1} (1-\alpha_j) \right)^{-1}. \qquad [3.1]$$

Since $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$, this is equivalent to the following (linear) recursion for the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$:

$$\psi_\ell = \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i \right) \psi_{\ell-1} - \sum_{i=1}^{\min\{\ell-1,n\}} \varphi_i \psi_{\ell-1-i}, \qquad [3.2]$$

starting from $\psi_0 \equiv 1$. The stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$ is proportional to the sequence of survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$, so [3.2] implies that this age distribution is generated by an identical linear recursion, starting from $\omega_0 = \alpha$:

$$\omega_\ell = \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i \right) \omega_{\ell-1} - \sum_{i=1}^{\min\{\ell-1,n\}} \varphi_i \omega_{\ell-i-1}. \qquad [3.3]$$

[14] The presence of transitory idiosyncratic shocks can change this conclusion, though.

The formal justification for the interpretations of the parameters α and $\{\varphi_i\}_{i=1}^n$ is provided below. The overall strength of the selection effect will be measured by the difference between the average age of existing prices at the time of adjustment and the average age conditional on no adjustment.

Proposition 3 *Suppose the economy is at the unique stationary age distribution of prices* $\{\omega_\ell\}_{\ell=0}^\infty$.

(i) *The average probability of price adjustment* $\sum_{\ell=1}^\infty \alpha_\ell \omega_{\ell-1}$ *and the fraction of newly set prices* ω_0 *are both equal to* α.

(ii) *The expected duration of a newly set price* $(\sum_{\ell=1}^\infty \ell \alpha_\ell \psi_{\ell-1})/(\sum_{\ell=1}^\infty \alpha_\ell \psi_{\ell-1})$ *is* α^{-1}.

(iii) *The difference between the average age of existing prices that are adjusted and the average age conditional on no adjustment is* $(\sum_{i=1}^n i\varphi_i)\,\alpha^{-1}(1-\alpha)^{-1}$.

(iv) *The average age of prices in use* $\sum_{\ell=0}^\infty \ell\omega_\ell$ *is equal to* $(1-\sum_{i=1}^n i\varphi_i)\,\alpha^{-1}-1$.

(v) *The probability* α_1 *is strictly decreasing in all* φ_i, *and the probability* $\alpha_\infty \equiv \lim_{\ell\to\infty}\alpha_\ell$ *is strictly increasing in all* φ_i.

(vi) *The hazard function is flat* ($\alpha_\ell = \alpha$ *for all* ℓ) *if and only if* $\varphi_i = 0$ *for all* i.

(vii) *The hazard function is (weakly) upward sloping everywhere* ($\alpha_{\ell+1} \geq \alpha_\ell$ *for all* ℓ) *if* $\varphi_i \geq 0$ *for all* i.

(viii) *The hazard function is (weakly) downward sloping everywhere* ($\alpha_{\ell+1} \leq \alpha_\ell$ *for all* ℓ) *only if* $\varphi_i \leq 0$ *for all* i.

Proof See appendix A.4. ■

These results demonstrate how selection of the firms that get to change price is controlled by the parameters $\{\varphi_i\}_{i=1}^n$, with α simply measuring the overall extent of price flexibility. An increase in any φ_i raises the average age of existing prices in adjusting firms relative to non-adjusters. Equivalently, a rise in any one of these parameters pivots the hazard function around a unchanging average probability of price adjustment α by lowering the hazard function at short durations and raising it at long durations. Lowering any φ_i has the opposite effect. Multiple selection-effect parameters may be required because the hazard function need not be monotonic in general.

This ability to capture the direction and strength of the duration selection effect with a sequence of parameters $\{\varphi_i\}_{i=1}^n$ is very general in that essentially all hazard functions can be generated by a recursion of the form [3.1] as $n \to \infty$. Even though an infinite-dimensional parameter space is required to capture all possible hazard functions, the magnitude of parameter φ_i necessarily tends to zero as i increases for any hazard function implying convergence to the stationary age distribution.

Proposition 4 *A hazard function* $\{\alpha_\ell\}_{\ell=1}^\infty$ *can be exactly generated by the recursion [3.1] for some parameters* α *and* $\{\varphi_i\}_{i=1}^n$ *with* $n\to\infty$ *and* $\sum_{i=1}^\infty |\varphi_i| < \infty$ *if and only if it implies convergence to the stationary age distribution* $\{\omega_\ell\}_{\ell=0}^\infty$.

Proof See appendix A.5. ■

Hence, any hazard function implying convergence to the stationary age distribution can be arbitrarily well approximated by [3.1] with a sufficiently large, but finite, number of parameters α and $\{\varphi_i\}_{i=1}^n$ because it is known that $\varphi_i \to 0$ as $i \to \infty$.

An arbitrary choice of parameters α and $\{\varphi_i\}_{i=1}^n$ does not necessarily yield a well-defined sequence of probabilities $\{\alpha_\ell\}_{\ell=1}^\infty$ using [3.1]. The following result presents some restrictions on the parameters that need to be satisfied.

Proposition 5 *(i) Parameters consistent with a well-defined hazard function $\{\alpha_\ell\}_{\ell=1}^\infty$ must satisfy $0 \leq \alpha \leq 1$ and $-(1-\alpha) \leq \sum_{j=i}^n \varphi_j \leq \alpha$ for all $i = 1, \ldots, n$.*

(ii) In the case $n = 1$, the restrictions $0 \leq \alpha \leq 1$ and $-\sqrt{\alpha}(1-\sqrt{\alpha}) \leq \varphi \leq \min\{\alpha, (1-\sqrt{\alpha})^2\}$ on α and $\varphi \equiv \varphi_1$ are necessary and sufficient for the hazard function to be well defined.

Proof See appendix A.6. ■

4 Inflation dynamics

Having established a parameterization of the hazard function that ties the selection effect to $\{\varphi_i\}_{i=1}^n$, the next step is to understand how the selection effect changes inflation dynamics relative to the New Keynesian Phillips curve, where no such selection effect is present. The analysis is facilitated by the fact that the hazard function recursion [3.1] with a finite number of parameters generates a "hybrid" Phillips curve of the kind [2.7] presented in section 2 with a *finite* number of leads and lags of inflation.

To see this, note that equation [3.2] implies the reset price equation [2.4] can be written equivalently as

$$r_t = \beta\left(1 - \alpha + \sum_{i=1}^n \varphi_i\right)\mathbb{E}_t r_{t+1} - \sum_{\ell=1}^n \beta^{\ell+1}\varphi_\ell \mathbb{E}_t r_{t+1+\ell} + \left(1 - \beta(1-\alpha) - \beta\sum_{i=1}^n (1-\beta^i)\varphi_i\right)p_t^*. \quad [4.1]$$

Similarly, given [3.3], the equation for the price level in [2.5] becomes

$$p_t = \left(1 - \alpha + \sum_{i=1}^n \varphi_i\right)p_{t-1} - \sum_{\ell=1}^n \varphi_\ell p_{t-1-\ell} + \alpha r_t. \quad [4.2]$$

The parameterization [3.1] of the hazard function thus justifies finite-order *recursive* equations for the reset price r_t and the price level p_t if the number of parameters is finite. In essence, this recursive structure is what gives the New Keynesian Phillips curve its simple form. But Calvo pricing specifically requires *first-order* recursions in [4.1] and [4.2]. The presence of a selection effect is equivalent to the recursions being of a higher order, but this does not mean the highly convenient recursive structure cannot be exploited.

Equations [2.3], [4.1] and [4.2] are combined to derive the Phillips curve. The result below presents its properties.

Proposition 6 *(i) The Phillips curve relating inflation π_t to real marginal cost x_t is*

$$\pi_t = \sum_{\ell=1}^n \lambda_\ell \pi_{t-\ell} + \sum_{\ell=1}^{n+1} \xi_\ell \mathbb{E}_t \pi_{t+\ell} + \nu\kappa x_t. \quad [4.3]$$

(ii) *The current inflation response to current real marginal cost is always positive ($\kappa > 0$).*

(iii) *Current inflation always depends positively on the expectation of next period's inflation rate ($\xi_1 > 0$).*

(iv) *The coefficient λ_ℓ on lagged inflation ℓ periods ago is a weighted sum (with non-negative weights) of the selection-effect parameters $\varphi_\ell, \ldots, \varphi_n$.*

(v) *The coefficient ξ_ℓ on future inflation ℓ periods in the future is the negative of a weighted sum (with non-negative weights) of the selection-effect parameters $\varphi_{\ell-1}, \ldots, \varphi_n$ for $\ell \geq 2$.*

(vi) *The coefficients on lagged and future inflation sum to unity when weighted by the discount factor β: $\sum_{\ell=1}^{n} \beta^\ell \lambda_\ell + \sum_{\ell=1}^{n+1} \beta^{-\ell} \xi_\ell = 1$.*

(vii) *The following cross-coefficient restrictions are satisfied for all hazard functions: $\xi_1 = \beta + (1-\beta)\sum_{i=1}^{n} \beta^i \lambda_i$ and $\xi_\ell = -\left\{\beta^\ell \lambda_{\ell-1} - (1-\beta)\sum_{i=\ell}^{n} \beta^i \lambda_i\right\}$ for $\ell = 2, \ldots, n+1$.*

Proof See appendix A.7. ■

First, the general Phillips curve [4.3] always shares two features of the NKPC: the positive dependence of current inflation on current real marginal cost and on expectations of inflation one period ahead. The existence of a selection effect shows up in the presence of lags of inflation and extra leads of inflation. A positive selection effect (firms being more likely to adjust prices set long ago than those set recently) implies a positive dependence of current inflation on past inflation, while a negative selection effect (recently set prices being more likely to be changed than older prices) implies a negative relationship between current and past inflation. A positive selection effect thus provides a microfoundation for intrinsic inflation persistence.

To understand the intuition for this finding, suppose that at time $t = 1$, a temporary shock to real marginal cost x_1 occurs that raises the inflation rate π_1 by 1%. Normalize the time $t = 0$ price level to $\mathsf{p}_0 = 0$ and suppose there is no price dispersion at $t = 0$ for simplicity. Given that only a fraction α $(0 < \alpha < 1)$ of firms changes price at time t, it follows that the reset price of those firms that adjusted is $\mathsf{r}_1 = \alpha^{-1} > 1$. Now consider the conjecture that the price level p_t is expected to remain 1% higher permanently. In the absence of further shocks to real marginal cost, this belief implies that $\mathsf{r}_t = 1$ is the best reset price for any firm from period $t = 2$ onwards. This leads to two effects on actual inflation: a "catch-up" effect of firms that did not initially respond to the shock now increasing their prices by 1% in line with the now-higher general price level; and a "roll-back" effect of those firms that did raise price initially now cutting back their prices by $(\alpha^{-1} - 1)$% to bring them into line with the general price level.

After period $t = 1$, the sizes of the groups of firms that want respectively to "catch up" or to "roll back" are proportional to $1 - \alpha$ and α. When firms from the "catch-up" group change price, their contribution to aggregate inflation (relative to their size) is $(1-\alpha)$%. Price changes from firms in the "roll-back" group contribute $-\alpha \times (\alpha^{-1} - 1) = -(1-\alpha)$% to inflation (again, relative to their size). Thus, if firms in the "catch-up" group are as likely to change price as those in the "roll-back" group then the overall effect on inflation is zero and the conjectured belief of no inflation persistence is confirmed as the rational expectations equilibrium. If there is a positive selection effect then "catch-up" is more likely than "roll-back", so actual inflation would continue to be positive after

$t = 1$, even though the shock is gone. This means that persistence is present in inflation even though the fundamentals explaining inflation themselves exhibit no persistence.

The impact of the selection effect on the coefficients of future inflation in [4.3] is the reverse of the finding for the lags of inflation. A positive selection effect implies negative coefficients of expected inflation more than one period in the future, reducing the overall importance of expected inflation in determining current inflation. These coefficients on the future inflation are essentially the mirror images of those on lagged inflation, and it is possible to deduce the former from the latter (together with knowledge of the discount factor). Note that inflation can never be entirely backward looking: there is always a positive dependence on inflation expectations one period ahead. A positive dependence on past inflation must come at the expense of the overall importance of future inflation because the coefficients on all inflation terms past and future must sum to one (when weighted by the discount factor).

These findings can be illustrated with a simple example where there is only one selection-effect parameter $\varphi \equiv \varphi_1$ ($n = 1$). In this case, the Phillips curve [4.3] reduces to

$$\pi_t = \lambda\pi_{t-1} + \beta(1 + (1 - \beta)\lambda)\mathbb{E}_t\pi_{t+1} - \beta^2\lambda\mathbb{E}_t\pi_{t+2} + \nu\kappa x_t, \qquad [4.4]$$

using the link between the future and past inflation coefficients, and where λ and κ are given by

$$\lambda = \frac{\varphi}{1 - \alpha + \beta\varphi(1 - \alpha + \varphi)}, \quad \text{and} \quad \kappa = \frac{\alpha(1 - \beta(1 - \alpha) - \beta(1 - \beta)\varphi)}{1 - \alpha + \beta\varphi(1 - \alpha + \varphi)}. \qquad [4.5]$$

The coefficient of lagged inflation λ has the same sign as that the selection effect parameter φ and is strictly increasing in its magnitude.[15] The absolute value of λ is strictly increasing in α; this is because the impact of the selection effect in making current inflation depend on past inflation is greater when prices are more flexible on average. Unsurprisingly, the component of the slope of the Phillips curve due to nominal rigidities κ (with ν depending on real rigidities) is strictly increasing in the average flexibility of prices α. The coefficient κ is also strictly decreasing in φ, so a positive selection effect reduces the overall slope of the Phillips curve. The coefficient on expected future inflation two periods ahead has the opposite sign to that of the selection effect parameter, though the sum of all coefficients on future inflation is always positive.

5 Estimating the hazard function

This section shows how the hazard function consistent with aggregate inflation dynamics can be estimated easily in a manner similar to how the average frequency of price adjustment is derived from estimates of the New Keynesian Phillips curve (a procedure which is a special case of the one developed in this paper). The method is implemented here using the simplest econometric techniques that have been applied in empirical work on the New Keynesian Phillips curve, that

[15] These results of course restrict α and φ to the range consistent with a well-defined hazard function as characterized by Proposition 5. The expressions for λ and κ are special cases of those derived in the proof of Proposition 6.

is, single-equation estimates that do not embed the Phillips curve in a complete structural model. The results are illustrative — recent work by Benati (2009) has estimated the model of this paper using more advanced Bayesian methods. Here, the econometrics will follow Galí and Gertler (1999) wherever possible.

Suppose that the Phillips curve [4.3] holds up to an additive i.i.d. measurement error υ_t. Assuming a Cobb-Douglas production function implies that the (log) labour share s_t is a proxy for the otherwise unobservable (log) real marginal cost x_t (both given as deviations from their steady-state values), as shown by Galí, Gertler and López-Salido (2001). Hence the Phillips curve can be written as:

$$\pi_t = \sum_{\ell=1}^{n} \lambda_\ell \pi_{t-\ell} + \sum_{\ell=1}^{n+1} \xi_\ell \mathbb{E}_t \pi_{t+\ell} + \nu\kappa \mathsf{s}_t + \upsilon_t.$$

Now define the ℓ-step ahead forecast errors of inflation $\Upsilon_{\ell,t} \equiv \pi_t - \mathbb{E}_{t-\ell}\pi_t$. Noting that $\mathbb{E}_t\pi_{t+\ell} = \pi_{t+\ell} - \Upsilon_{\ell,t+\ell}$, and making use of the expressions for ξ_ℓ in terms of $\{\lambda_\ell\}_{\ell=1}^n$ and β given in Proposition 6, it follows that the Phillips curve equation above becomes:

$$\{\pi_t - \beta\pi_{t+1}\} - \sum_{\ell=1}^{n} \lambda_\ell \left\{ \pi_{t-\ell} + (1-\beta)\beta^\ell \sum_{i=1}^{\ell} \pi_{t+i} - \beta^{\ell+1}\pi_{t+\ell+1} \right\} - \nu\kappa\mathsf{s}_t = \upsilon_t - \sum_{\ell=1}^{n+1} \xi_\ell \Upsilon_{\ell,t+\ell}.$$

Now take a vector of variables $\mathbf{z}_{t-1}$ in the information set of firms at time $t-1$. Rational expectations implies that these variables are uncorrelated with $\Upsilon_{\ell,t+\ell}$ for all $\ell \geq 0$, and they must also be uncorrelated with υ_t. Thus, taking the unconditional expectation of the equation above multiplied by the vector $\mathbf{z}_{t-1}$ yields the following moment conditions:

$$\mathbb{E}\left[\left\{ \{\pi_t - \beta\pi_{t+1}\} - \sum_{\ell=1}^{n} \lambda_\ell \left\{ \pi_{t-\ell} + (1-\beta)\beta^\ell \sum_{i=1}^{\ell} \pi_{t+i} - \beta^{\ell+1}\pi_{t+\ell+1} \right\} - \nu\kappa\mathsf{s}_t \right\} \mathbf{z}_{t-1} \right] = \mathbf{0}. \quad [5.1]$$

There are $n+3$ parameters to estimate (β, ν, α, $\{\varphi_i\}_{i=1}^n$ — treating real rigidity ν from [2.3] as a single parameter). Proposition 6 implies that the coefficients $\{\lambda_\ell\}_{\ell=1}^n$ and κ are functions of β, α and $\{\varphi_i\}_{i=1}^n$. However, it is clear from the form of the moment conditions in [5.1] that even with more than $n+3$ instruments in $\mathbf{z}_{t-1}$, at most $n+2$ parameters can be identified. Thus, as in Galí, Gertler and López-Salido (2001), the real rigidity parameter ν must be calibrated.

Given a Cobb-Douglas production function and monopolistic competition modelled using a CES aggregator, it follows that $\nu = ms/(1-s)$, where s is the average labour share and m is the average net markup of price on marginal cost. Calibrating the former to 67% and the latter to 10% yields $\nu = 0.2$.[16]

With this calibration, all the remaining $n+2$ parameters β, α and $\{\varphi_i\}_{i=1}^n$ can be estimated by the Generalized Method of Moments (GMM) if at least $n+2$ instruments can be found.[17] Extra in-

[16] This calibration follows Gertler and Leahy (2008). The original study by Galí and Gertler (1999) imposed a linear production function, which implies $\nu = 1$.

[17] In small samples, the normalization of the moment conditions [5.1] can affect the results. Following Galí and Gertler (1999), equation [5.1] is multiplied by a function of the parameters that ensures all the coefficients are bounded when the underlying parameters are restricted to a bounded set. This requires multiplying [5.1] by the expression for

struments provide over-identifying restrictions. The cross-coefficient restrictions from Proposition 6 are imposed, otherwise there would be multiple (and generally inconsistent) ways of recovering the hazard function from different combinations of Phillips-curve coefficients.[18] What is not imposed are the restrictions on the parameters α and $\{\varphi_i\}_{i=1}^{n}$ ensuring the hazard function is well defined (see Proposition 5). After estimation, it can be checked where there is any statistically significant deviation from a well-defined hazard function.[19]

Quarterly U.S. data on inflation and the labour share from 1960:Q1 to 2003:Q4 are used to estimate the hazard function. The GMM estimation procedure requires that instruments be found for the current and future endogenous variables appearing in the Phillips curve. The lags of the following variables were selected for this role in addition to lags of inflation and the labour share themselves: the spread between ten-year Treasury Bond and the three-month Treasury Bill yields, quadratically detrended log real GDP, wage inflation, and commodity-price inflation.[20] These are very similar to the instruments used in the original Galí and Gertler (1999) study of the NKPC.[21]

It is necessary to choose the number n of hazard-function slope (selection effect) parameters to estimate. Three cases are considered: the case of a flat hazard function ($n = 0$), equivalent to the New Keynesian Phillips curve; hazard functions consistent with one lag of inflation in the Phillips curve ($n = 1$); and hazard functions consistent with four lags of inflation ($n = 4$).[22] The estimation results are presented in Table 1.[23] In all cases, the estimated quarterly frequency of price adjustment α is approximately between 20% and 25%. The estimated discount factor β is less than one, but not significantly so. The first hazard-function slope parameter φ_1 is significantly positive in both the $n = 1$ and $n = 4$ specifications. When $n = 4$, there is a mixture of (significantly) positive and negative estimated slope parameters. The implied hazard functions are shown in Figure 1 and Figure 2.

The hazard function in the case $n = 1$ is upward sloping everywhere, rising from around zero for newly set prices to a probability of approximately 0.4 after one year, and remaining largely flat after that. In the case of $n = 4$, the hazard function is no longer monotonic. It also starts close to

θ_0 from appendix A.7.

[18]These restrictions could in principle be tested.

[19]The restrictions needed for this are equivalent to inequality constraints on the parameters, so are more difficult to impose at the estimation stage.

[20]The source of the data is the Federal Reserve Economic Data (FRED) database, available online at `http://research.stlouisfed.org/fred2`. Inflation is measured by the annualized percentage change in the GDP deflator between consecutive quarters. The labour share is given by unit labour costs in the business sector divided by the GDP deflator. Wage inflation is the annualized percentage change in compensation per hour in the business sector. Commodity-price inflation is measured by the percentage change between consecutive quarters of a commodities futures-price index. All variables are expressed as deviations from their averages over the sample period.

[21]Based on their predictive power for future inflation, six lags of inflation and commodity-price inflation were selected as instruments, together with two lags of each of the other variables.

[22]No extra parameters were found to be statistically significant for $n > 4$.

[23]The GMM estimator applied in this paper uses a four-lag Newey-West estimator of the optimal weighting matrix for the moment conditions. For each weighting matrix, the numerical minimization algorithm for the criterion function is iterated until convergence because the coefficients in the moment conditions are non-linear functions of the parameters. The resulting estimates are then used to update the weighting matrix, and the process is repeated until the weighting matrix converges itself. Robust standard errors of the parameter estimates are also obtained using a four-lag Newey-West estimator of the variance-covariance matrix. Details of these procedures are provided in Mátyás (1999).

Table 1: *Estimated parameter values*

n	α	φ_1	φ_2	φ_3	φ_4	β	J-stat[†]
0	0.259**					0.973**	16.093
	(0.048)					(0.026)	[0.651]
1	0.245**	0.258**				0.920**	10.983
	(0.049)	(0.077)				(0.069)	[0.895]
4	0.195**	0.182**	−0.026	−0.208**	0.214**	0.874**	5.932
	(0.055)	(0.081)	(0.070)	(0.087)	(0.054)	(0.074)	[0.981]

Notes: Estimation of the parameters is by GMM using U.S. quarterly data 1960:Q1–2003:Q4 and moment conditions [5.1]. The estimators of the parameters and the GMM weighting matrix are sequentially iterated until convergence. A four-lag Newey-West estimator of the optimal weighting matrix and the standard errors is used. Standard errors are given in parentheses, and are calculated using the delta method for non-linear functions of the estimated parameters.

* Statistically significant at the 10% level.

** Statistically significant at the 5% level.

† This is the Hansen test of over-identifying moment conditions. The p-value is in brackets.

zero, rises during the first year, before diminishing toward the end of the year. It then rises sharply after the first year is over, and repeats a similar pattern toward the end of the second year.[24] The hazard function finishes much higher (around 0.4) than it begins (around 0), so it is still upward sloping overall. Note that the point estimate of the probability at the four-quarter duration is not well defined, but this deviation is not statistically significant.

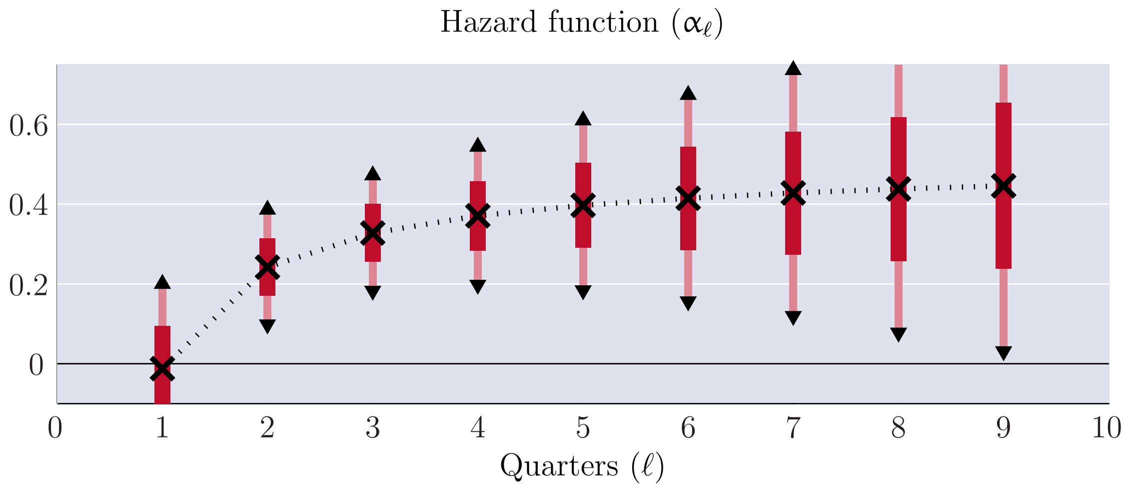
The coefficients of the implied Phillips curves are shown in Table 2. Inflation one quarter ago has a significantly positive coefficient in both the $n = 1$ and $n = 4$ specifications, while for $n = 4$ there is also a significantly positive coefficient on inflation four quarters ago. This latter finding reflects the dip and subsequent rise of the hazard function around durations of one year.

6 Conclusions

This paper has shown that it is not necessary to appeal to backward-looking pricing rules to explain why inflation can display intrinsic persistence. Relaxing the commonly used Calvo (1983) assumption of random selection of which prices are changed, and allowing for the probability of adjustment to depend on the duration of price stickiness, implies that the Phillips curve takes on a "hybrid" form, including both past and expected future inflation. In general, the sign of the coefficients

[24] The model is set up in discrete time and the data used for the estimation are quarterly, so the results cannot reveal whether a change in the hazard function occurs at the beginning, middle, or end of any given quarter. This means that it is hard to judge whether the rise in the hazard function after the first and second years is consistent with the timing of the observed spikes in microeconometric hazard functions at 12-month intervals.

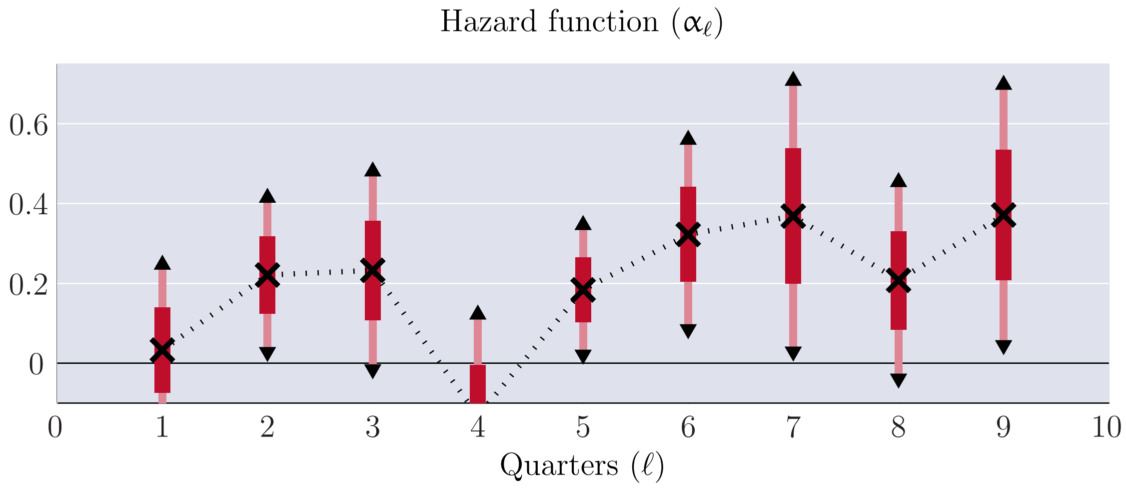
Figure 1: *Hazard function for first-order model*



Notes: The hazard function is calculated from [3.1] using the estimated parameters from Table 1. The crosses joined by the dotted lines are the point estimates, the darkly shaded thick bars are one-standard-deviation confidence intervals; the lightly shaded thin bars are two-standard-deviation confidence intervals. Confidence bands are calculated using the delta method.

on lagged inflation can be either positive or negative depending on whether the selection effect is positive or negative, that is, whether the probability of adjustment increases or decreases with the duration of a price spell. If the hazard function is upward sloping then lagged inflation rates receive positive coefficients; if downward-sloping then the coefficients are negative. Empirical estimates suggest that there is a well-defined hazard function consistent with U.S. inflation dynamics. This hazard function is not monotonic, but is clearly upward sloping on average.

Figure 2: *Hazard function for fourth-order model*



Notes: See notes to Figure 1.

Table 2: *Coefficients of Phillips curves*

n	π_{t-4}	π_{t-3}	π_{t-2}	π_{t-1}	$\mathbb{E}_t\pi_{t+1}$	$\mathbb{E}_t\pi_{t+2}$	$\mathbb{E}_t\pi_{t+3}$	$\mathbb{E}_t\pi_{t+4}$	$\mathbb{E}_t\pi_{t+5}$	x_t
0					0.973**					0.080**
					(0.026)					(0.032)
1				0.259**	0.939**	−0.219**				0.056**
				(0.052)	(0.052)	(0.040)				(0.028)
4	0.220**	−0.180**	0.005	0.182**	0.896**	−0.137**	−0.002	0.121**	−0.113**	0.044
	(0.058)	(0.080)	(0.065)	(0.067)	(0.062)	(0.059)	(0.047)	(0.059)	(0.049)	(0.028)

Notes: This table reports the coefficients of the Phillips curve [4.3] implied by the parameter estimates in Table 1. Standard errors are given in parentheses and are calculated using the delta method. See the notes to Table 1 for more details about the estimation method.

* Statistically significant at the 10% level.

** Statistically significant at the 5% level.

References

Álvarez, L. J., Burriel, P. and Hernando, I. (2005), "Do decreasing hazard functions for price changes make any sense?", Working paper 461, European Central Bank. 3

Bakhshi, H., Khan, H. and Rudolf, B. (2007), "The Phillips curve under state-dependent pricing", *Journal of Monetary Economics*, **54(8)**:2321–2345 (November). 4

Ball, L. and Romer, D. (1990), "Real rigidities and the non-neutrality of money", *Review of Economic Studies*, **57(2)**:183–203 (April). 5

Benati, L. (2009), "Are 'intrinsic inflation persistence' models structural in the sense of Lucas (1976)?", Working paper 1038, European Central Bank. 15

Calvo, G. A. (1983), "Staggered prices in a utility-maximizing framework", *Journal of Monetary Economics*, **12(3)**:383–398 (September). 1, 8, 17

Cavallo, A. (2009), "Scraped data and sticky prices: Frequency, hazards, and synchronization", Working paper, Harvard University. 3

Cecchetti, S. G. (1986), "The frequency of price adjustment: A study of the newsstand prices of magazines", *Journal of Econometrics*, **31(3)**:255–274 (April). 3

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005), "Nominal rigidities and the dynamic effects of a shock to monetary policy", *Journal of Political Economy*, **113(1)**:1–45 (February). 1, 8

COGLEY, T. AND SBORDONE, A. M. (2005), "A search for a structural Phillips curve", Staff report 203, Federal Reserve Bank of New York. 1

DANZIGER, L. (1999), "A dynamic economy with costly price adjustments", *American Economic Review*, **89(4)**:878–901 (September). 9

DOTSEY, M. (2002), "Pitfalls in interpreting tests of backward-looking pricing in New Keynesian models", *Federal Reserve Bank of Richmond Economic Quarterly*, **88(1)**:37–50 (Winter). 3

DOTSEY, M., KING, R. G. AND WOLMAN, A. L. (1999), "State dependent pricing and the general equilibrium dynamics of money and output", *Quarterly Journal of Economics*, **114(2)**:655–690 (May). 3

EICHENBAUM, M., JAIMOVICH, N. AND REBELO, S. (2008), "Reference prices and nominal rigidities", Working paper 13829, National Bureau of Economic Research. 3

FUHRER, J. AND MOORE, G. (1995), "Inflation persistence", *Quarterly Journal of Economics*, **109(1)**:197–230 (February). 1

FUHRER, J. C. (1997), "The (un)importance of forward-looking behavior in price specifications", *Journal of Money, Credit and Banking*, **29(3)**:338–350 (August). 3

——— (2010), "Inflation persistence", in B. M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics*, Elsevier, (forthcoming). 1

GALÍ, J. AND GERTLER, M. (1999), "Inflation dynamics: A structural econometric analysis", *Journal of Monetary Economics*, **44(2)**:195–222 (October). 1, 8, 15, 16

GALÍ, J., GERTLER, M. AND LÓPEZ-SALIDO, J. D. (2001), "European inflation dynamics", *European Economic Review*, **45(7)**:1237–1270 (June). 15

GAMELIN, T. W. (2001), *Complex Analysis*, Springer-Verlag, New York. 24

GERTLER, M. AND LEAHY, J. (2008), "A Phillips curve with an Ss foundation", *Journal of Political Economy*, **116(3)**:533–572. 9, 15

GOLOSOV, M. AND LUCAS, R. E., JR. (2007), "Menu costs and Phillips curves", *Journal of Political Economy*, **115(2)**:171–199. 9, 10

GOODFRIEND, M. AND KING, R. G. (1997), "The New Neoclassical synthesis and the role of monetary policy", *NBER Macroeconomics Annual*, **12**:231–283. 3

GÖTTE, L., MINSCH, R. AND TYRAN, J.-R. (2005), "Micro evidence on the adjustment of sticky-price goods: It's how often, not how much", Discussion paper 5364, Centre for Economic Policy Research. 3

Guerrieri, L. (2001), "Inflation dynamics", International finance discussion paper 715, Federal Reserve Board. 3

——— (2002), "The inflation persistence of staggered contracts", International finance discussion paper 734, Federal Reserve Board. 3

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press. 7

Ikeda, D. and Nishioka, S. (2007), "Price setting behavior and hazard functions: Evidence from Japanese CPI micro data", Working paper 07-E-19, Bank of Japan. 3

Jadresic, E. (1999), "Sticky prices: An empirical assessment of alternative models", Working paper 99/72, International Monetary Fund. 2

Khan, A., King, R. G. and Wolman, A. L. (2003), "Optimal monetary policy", *Review of Economic Studies*, **70(4)**:825–860 (October). 2

Klenow, P. J. and Malin, B. A. (2010), "Microeconomic evidence on price-setting", in B. M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics*, Elsevier, (forthcoming). 3

Laforte, J.-P. (2007), "Pricing models: A Bayesian DSGE approach for the U.S. economy", *Journal of Money, Credit and Banking*, **39(s1)**:127–154 (February). 2, 3

Mankiw, N. G. and Reis, R. (2002), "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve", *Quarterly Journal of Economics*, **117(4)**:1295–1328 (November). 3, 9

Mash, R. (2004), "Optimizing microfoundations for inflation persistence", Discussion paper 183, Oxford University. 3, 6

Mátyás, L. (1999), *Generalized Method of Moments Estimation*, Cambridge University Press, Cambridge. 16

Milani, F. (2005), "Adaptive learning and inflation persistence", Working paper, Princeton University. 1

Nakamura, E. and Steinsson, J. (2008), "Five facts about prices: A reevaluation of menu cost models", *Quarterly Journal of Economics*, **123(4)**:1415–1464 (November). 3, 6

Paloviita, M. (2004), "Inflation dynamics in the euro area and the role of expectations: Further results", Discussion paper 21, Bank of Finland. 1

Roberts, J. M. (1997), "Is inflation sticky?", *Journal of Monetary Economics*, **39(2)**:173–196 (July). 1

——— (2005), "How well does the New Keynesian sticky-price model fit the data?", *Contributions to Macroeconomics*, **5(1)**:Article 10. 1

Rudd, J. and Whelan, K. (2005), "New tests of the New-Keynesian Phillips curve", *Journal of Monetary Economics*, **52(6)**:1167–1181 (September). 1

Sheedy, K. D. (2007), "Robustly optimal monetary policy", Discussion paper 840, Centre for Economic Performance. 2

Smets, F. and Wouters, R. (2003), "An estimated dynamic stochastic general equilibrium model of the euro area", *Journal of the European Economic Association*, **1(5)**:1123–1175 (September). 1, 8

Taylor, J. B. (1980), "Aggregate dynamics and staggered contracts", *Journal of Political Economy*, **88(1)**:1–24 (February). 3, 6

Whelan, K. (2007), "Staggered price contracts and inflation persistence: Some general results", *International Economic Review*, **48(1)**:111–145 (February). 3, 6

Wolman, A. L. (1999), "Sticky prices, marginal cost, and the behavior of inflation", *Federal Reserve Bank of Richmond Economic Quarterly*, **85(4)**:29–48 (Fall). 2, 3, 4

Woodford, M. (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, New Jersey. 4, 8

——— (2007), "Interpreting inflation persistence: Comments on the conference on 'Quantitative Evidence on Price Determination'", *Journal of Money, Credit and Banking*, **39(s1)**:203–210 (February). 1

——— (2010), "Optimal monetary stabilization policy", in B. M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics*, Elsevier, (forthcoming). 2

Yao, F. (2009), "Non-constant hazard function and inflation dynamics", Discussion paper 2009-030, Humboldt Universität. 3, 6

A Technical appendix

A.1 Lemmas

Lemma 1 *Let* $\mathscr{D}_\rho \equiv \left\{ z \in \mathbb{C} \;\middle|\; |z| \leq \rho \right\}$ *be the closed disc of radius* ρ *and let* $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ *denote the* z*-transform of the sequence of survival probabilities* $\{\psi_\ell\}_{\ell=0}^{\infty}$*.*

(i) The function $\psi(z)$ *is analytic on* $\mathscr{D}_\rho$ *for some* $\rho > 1$*.*

(ii) A stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$ *consistent with the hazard function* $\{\alpha_\ell\}_{\ell=1}^{\infty}$ *is stable if and only if* $\psi(z)$ *has no roots in* $\mathscr{D}_\rho$ *for some* $\rho > 1$*.*

PROOF (i) Let $\alpha_\infty \equiv \lim_{\ell\to\infty} \alpha_\ell$ be the limiting value of the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$. It is assumed that $\alpha_\infty > 0$. Let ϖ be a number lying strictly between $(1-\alpha_\infty)$ and 1, which must satisfy $0 < \varpi < 1$. Given that $\alpha_\ell \to \alpha_\infty$ as $\ell \to \infty$, there must exist a value of $\jmath$ such that $(1-\alpha_\ell) < \varpi$ for all $\ell \geq \jmath$. Since $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$, this implies that $\psi_{\ell+1} \leq \varpi\psi_\ell$ for all $\ell \geq \jmath$ and hence $\psi_\ell \leq \varpi^{\ell-\jmath}\psi_\jmath$.

Now let ρ be any number strictly between 1 and ϖ^{-1}, which implies $0 < \varpi\rho < 1$ and hence that $\sum_{\ell=0}^{\infty} \varpi^\ell |z|^\ell \leq \sum_{\ell=0}^{\infty} (\varpi\rho)^\ell = (1-\varpi\rho)^{-1} < \infty$ for all $z \in \mathscr{D}_\rho$. By applying the triangle inequality to the power series $\psi(z)$, it follows that

$$\left|\sum_{\ell=0}^{\infty} \psi_\ell z^\ell\right| \leq \sum_{\ell=0}^{\infty} \psi_\ell |z|^\ell \leq \sum_{\ell=0}^{\jmath-1} \psi_\ell |z|^\ell + \psi_\jmath \sum_{\ell=\jmath}^{\infty} \varpi^{\ell-\jmath}|z|^\ell = \sum_{\ell=0}^{\jmath-1} \psi_\ell |z|^\ell + \psi_\jmath |z|^\jmath \sum_{\ell=0}^{\infty} \varpi^\ell |z|^\ell,$$

and hence $|\psi(z)| < \infty$ if $z \in \mathscr{D}_\rho$. Therefore, the power series $\psi(z)$ is analytic on $\mathscr{D}_\rho$ for some $\rho > 1$.

(ii) Let $\{\omega_\ell\}_{\ell=0}^{\infty}$ be a stationary age distribution satisfying $\sum_{\ell=0}^{\infty} \omega_\ell = 1$ and consistent with the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ so that $\omega_\ell = (1-\alpha_\ell)\omega_{\ell-1}$. Now let $\Delta_{\ell,t} \equiv \omega_{\ell,t} - \omega_\ell$ denote the sequence of deviations $\{\Delta_{\ell,t}\}_{\ell=0}^{\infty}$ from this stationary distribution at time t. As $\sum_{\ell=0}^{\infty} \omega_{\ell,t} = 1$, it must be the case that $\sum_{\ell=0}^{\infty} \Delta_{\ell,t} = 0$ and hence $\Delta_{0,t} = -\sum_{\ell=1}^{\infty} \Delta_{\ell,t}$. Thus considering the sequence of deviations $\{\Delta_{\ell,t}\}_{\ell=1}^{\infty}$ is sufficient to know the behaviour of the whole sequence $\{\Delta_{\ell,t}\}_{\ell=0}^{\infty}$.

The laws of motion for the age distribution $\{\omega_{\ell,t}\}_{\ell=0}^{\infty}$ require that $\omega_{\ell,t} = (1-\alpha_\ell)\omega_{\ell-1,t-1}$ for all $\ell \geq 1$. These imply laws of motion for the deviations $\Delta_{\ell,t}$:

$$\Delta_{1,t} = -(1-\alpha_1)\sum_{\ell=1}^{\infty} \Delta_{\ell,t-1}, \quad \text{and} \quad \Delta_{\ell,t} = (1-\alpha_\ell)\Delta_{\ell-1,t-1} \quad \text{for} \quad \ell = 2,3,\ldots, \qquad [A.1.1]$$

using the earlier formula for $\Delta_{0,t-1}$ in terms of the sequence $\{\Delta_{\ell,t-1}\}_{\ell=1}^{\infty}$.

The equations in [A.1.1] define a linear transformation of the sequence $\{\Delta_{\ell,t}\}_{\ell=1}^{\infty}$. Suppose ζ is an eigenvalue of this linear transformation, with the sequence $\{\mathrm{v}_\ell\}_{\ell=1}^{\infty}$ being the corresponding eigenvector. The eigenvalue-eigenvector pair is characterized by

$$\zeta\mathrm{v}_1 = -(1-\alpha_1)\sum_{\ell=1}^{\infty} \mathrm{v}_\ell, \quad \text{and} \quad \zeta\mathrm{v}_\ell = (1-\alpha_\ell)\mathrm{v}_{\ell-1} \quad \text{for} \quad \ell = 2,3,\ldots. \qquad [A.1.2]$$

The stability of the stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$ is equivalent to all eigenvalues of the linear transformation having modulus less than one.

For a non-zero eigenvalue ζ, note that the equations in [A.1.2] imply $\mathrm{v}_1 \neq 0$, otherwise all elements of the sequence $\{\mathrm{v}_\ell\}_{\ell=1}^{\infty}$ would be zero, which would mean that it could not be an eigenvector (which must be non-zero). Applying [A.1.2] recursively yields

$$(1-\alpha_1)\mathrm{v}_\ell = \zeta^{-(\ell-1)}\left\{\prod_{\jmath=1}^{\ell}(1-\alpha_\jmath)\right\}\mathrm{v}_1 \quad \text{for} \quad \ell = 2,3,\ldots,$$

and hence $(1-\alpha_1)\mathrm{v}_\ell = \zeta^{-(\ell-1)}\psi_\ell\mathrm{v}_1$ using the definition of the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$. Substitution into the remaining equation from [A.1.2] implies

$$\left\{\sum_{\ell=0}^{\infty} \psi_\ell \zeta^{-\ell}\right\}\mathrm{v}_1 = 0,$$

which together with $\mathrm{v}_1 \neq 0$ requires that $\psi(\zeta^{-1}) = 0$. Thus, any eigenvalue ζ of the linear transformation from $\{\Delta_{\ell,t}\}_{\ell=1}^{\infty}$ to $\{\Delta_{\ell,t+1}\}_{\ell=1}^{\infty}$ is either zero, or its reciprocal ζ^{-1} is a root of the equation $\psi(z) = 0$. Similarly, the reciprocal of any root of $\psi(z) = 0$ will be an eigenvalue of the linear transformation.

If there is a $\rho > 1$ such that $\psi(z) = 0$ has no roots on $\mathscr{D}_\rho$ then all eigenvalues ζ must have modulus less than one. Conversely, note that $\mathscr{D}_\rho$ is a compact set for any fixed ρ. If this ρ is no more than the threshold found in part (i) then $\psi(z)$ is an analytic function on $\mathscr{D}_\rho$, so it has a finite number of roots in this set. Hence if all eigenvalues ζ have modulus less than one then there exists a minimum value of $|\zeta^{-1}|$, which is greater than one. It follows that there exists a $\rho > 1$ such that $\psi(z) = 0$ has no roots on $\mathscr{D}_\rho$. This completes the proof. ■

Lemma 2 *Suppose $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ is a power series with coefficients satisfying $\psi_0 = 1$ and $0 \leq \psi_{\ell+1} \leq (1-\underline{\alpha})\psi_\ell$ for all $\ell \geq 0$ and for some $0 < \underline{\alpha} \leq 1$. Then there exists a $\rho > 1$ such that the equation $\psi(z) = 0$ has no roots in the set $\mathscr{D}_\rho \equiv \left\{ z \in \mathbb{C} \,\middle|\, |z| \leq \rho \right\}$.*

Proof Let ϖ be a number lying strictly between $(1-\underline{\alpha})$ and 1, which must satisfy $0 < \varpi < 1$. Since $\psi_{\ell+1} = (1-\alpha_{\ell+1})\psi_\ell$, the definition of ϖ then implies that $\psi_{\ell+1} \leq \varpi\psi_\ell$ for all $\ell \geq 0$. Now let ρ be any number strictly between 1 and the minimum of ϖ^{-1} and the radius of the disc on which $\psi(z)$ is analytic (greater than one), as established by Lemma 1.

Construct a new function $\mathfrak{F}(z) \equiv (1-\varpi z)\psi(z)$, which inherits the property that it is analytic on $\mathscr{D}_\rho$ from $\psi(z)$ using Lemma 1. Using the definition of $\psi(z)$ and collecting terms in common powers of z:

$$\mathfrak{F}(z) = 1 - \sum_{\ell=1}^{\infty} (\varpi\psi_{\ell-1} - \psi_\ell) z^\ell.$$

The function can be written as $\mathfrak{F}(z) = \mathfrak{F}_0(z) + \mathfrak{F}_1(z)$, where $\mathfrak{F}_0(z) \equiv 1$ and $\mathfrak{F}_1(z) \equiv -\sum_{\ell=1}^{\infty}(\varpi\psi_{\ell-1} - \psi_\ell)z^\ell$ are defined. The modulus of $\mathfrak{F}_1(z)$ satisfies

$$|\mathfrak{F}_1(z)| \leq \sum_{\ell=1}^{\infty} |\varpi\psi_{\ell-1} - \psi_\ell||z|^\ell = \sum_{\ell=1}^{\infty} (\varpi\psi_{\ell-1} - \psi_\ell)|z|^\ell,$$

using the triangle inequality and the positivity of the coefficient of $|z|^\ell$. Now take any $z \in \mathscr{D}_\rho$. Since $|z|^\ell \leq \rho^\ell$, it follows that

$$\sum_{\ell=1}^{\infty} (\varpi\psi_{\ell-1} - \psi_\ell)|z|^\ell \leq \sum_{\ell=1}^{\infty} (\varpi\psi_{\ell-1} - \psi_\ell)\rho^\ell = \varpi\rho - (1-\varpi\rho)\sum_{\ell=1}^{\infty} \psi_\ell \rho^\ell \leq \varpi\rho,$$

by collecting common terms in ψ_ℓ and using the non-negativity of $\{\psi_\ell\}_{\ell=0}^{\infty}$ together with $\psi_0 = 1$ and $0 < \varpi\rho < 1$. Combining the equations above yields $|\mathfrak{F}_1(z)| \leq \varpi\rho$, and hence $|\mathfrak{F}_1(z)| < |\mathfrak{F}_0(z)|$ for all $z \in \mathscr{D}_\rho$, since $|\mathfrak{F}_0(z)| = 1$.

As a constant function, $\mathfrak{F}_0(z)$ must be analytic, and consequently $\mathfrak{F}_1(z)$ inherits this property from $\mathfrak{F}(z)$. Since $\mathfrak{F}(z) = \mathfrak{F}_0(z) + \mathfrak{F}_1(z)$, Rouché's Theorem[25] implies that $\mathfrak{F}(z)$ and $\mathfrak{F}_0(z)$ have the same number of zeros on $\mathscr{D}_\rho$. Since $\mathfrak{F}_0(z)$ clearly has no zeros on this set, neither has $\mathfrak{F}(z)$. Because its definition ensures that $\mathfrak{F}(z)$ inherits any roots of $\psi(z) = 0$, this precludes $\psi(z)$ having a zero in $\mathscr{D}_\rho$ as well. This completes the proof. ■

Lemma 3 *The sequence of recursive parameters $\{\varphi_i\}_{i=1}^{\infty}$ generating the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ using [3.1] can be written as*

$$\varphi_i = (-1)^i \sum_{(\jmath_1,\ldots,\jmath_{i+1}) \in \mathscr{C}_{i+1}} \prod_{\ell=1}^{i+1} \sigma_{\jmath_\ell}, \qquad \text{[A.1.3]}$$

where the sequence $\{\sigma_\ell\}_{\ell=1}^{\infty}$ is defined by $\sigma_1 \equiv -(1-\alpha_1)$ and $\sigma_\ell \equiv \alpha_\ell - \alpha_{\ell-1}$, and where $\{\mathscr{C}_i\}_{i=2}^{\infty}$ is a sequence of sets $\mathscr{C}_i$, with each $\mathscr{C}_i$ being a subset of the set of sequences

$$\mathscr{P}_i \equiv \left\{ (\jmath_1, \ldots, \jmath_i) \in \mathbb{N}^i \,\middle|\, 1 \leq \jmath_\ell \leq \ell \right\}. \qquad \text{[A.1.4]}$$

[25] See any text on complex analysis, such as Gamelin (2001), for further details about the theorem.

Proof Define a sequence $\{\phi_i\}_{i=1}^{\infty}$ with $\phi_1 \equiv 1-\alpha$ and $\phi_i \equiv -\varphi_{i-1}$ for $i \geq 2$. With these definitions, the recursion [3.2] for the survival function $\{\psi_\ell\}_{\ell=0}^{\infty}$ reduces to:

$$\psi_\ell = \sum_{i=1}^{\ell} \phi_i \psi_{\ell-i},$$

with initial condition $\psi_0 = 1$. Using the initial condition, the order of the recursion can be reversed to yield

$$\phi_i = \psi_i - \sum_{j=1}^{i-1} \psi_{i-j}\phi_j. \qquad \text{[A.1.5]}$$

The definition of the survival probabilities means that $\psi_i = \prod_{\ell=1}^{i}(1-\alpha_\ell)$, and the definition of the sequence $\{\sigma_\ell\}_{\ell=1}^{\infty}$ in the statement of the Lemma implies $1-\alpha_\ell = \sum_{j=1}^{\ell}(-\sigma_j)$. It follows that

$$\psi_i = \prod_{\ell=1}^{i}\sum_{j=1}^{\ell}(-\sigma_j) = (-1)^i \sum_{\jmath_1=1}^{1} \cdots \sum_{\jmath_i=1}^{i} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell},$$

where the order of summation and multiplication is reversed in the final expression for ψ_i. Note that the definition of the set $\mathscr{P}_i$ of sequences $(\jmath_1, \ldots, \jmath_i)$ in [A.1.4] implies that ψ_i can be written as a sum of products $\prod_{\ell=1}^{i} \sigma_{\jmath_\ell}$ over all sequences in the set $\mathscr{P}_i$:

$$\psi_i = (-1)^i \sum_{(\jmath_1,\ldots,\jmath_i)\in\mathscr{P}_i} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell}. \qquad \text{[A.1.6]}$$

Now let $\mathscr{C}_1 \equiv \mathscr{P}_1 \equiv \{\ (1)\ \}$, where the expression for $\mathscr{P}_1$ comes from [A.1.4], and define the sets $\mathscr{C}_i$ in the sequence $\{\mathscr{C}_i\}_{i=2}^{\infty}$ with the recursion

$$\mathscr{C}_i \equiv \mathscr{P}_i \backslash \left(\bigcup_{j=1}^{i-1} (\mathscr{C}_j \times \mathscr{P}_{i-j})\right), \qquad \text{[A.1.7]}$$

in terms of the sequence $\{\mathscr{P}_i\}_{i=1}^{\infty}$ specified in [A.1.4]. Observe that $\mathscr{C}_i \subseteq \mathscr{P}_i$ is well defined if $\mathscr{C}_j \subseteq \mathscr{P}_j$ for all $j = 1, \ldots, i-1$ because $(\jmath_1, \ldots, \jmath_{i-j}) \in \mathscr{P}_{i-j}$ implies $\jmath_\ell \leq \ell + j$. Since $\mathscr{C}_1 \subseteq \mathscr{P}_1$ by definition, the claim that $\mathscr{C}_i \subseteq \mathscr{P}_i$ for all i follows by induction.

Now consider the following claim about the sequence of sets $\{\mathscr{C}_i\}_{i=1}^{\infty}$ defined by [A.1.7]:

$$(\mathscr{C}_j \times \mathscr{P}_{i-j}) \cap (\mathscr{C}_k \times \mathscr{P}_{i-k}) = \emptyset, \quad \text{for all } i, j, k \in \mathbb{N} \text{ with } j, k < i \ , \ j \neq k. \qquad \text{[A.1.8]}$$

Suppose for contradiction that $(\mathscr{C}_j \times \mathscr{P}_{i-j}) \cap (\mathscr{C}_k \times \mathscr{P}_{i-k}) \neq \emptyset$, and without loss of generality take $j > k$. Hence there is a sequence $(\jmath_1, \ldots, \jmath_i) \in \mathbb{N}^i$ such that $(\jmath_1, \ldots, \jmath_j) \in \mathscr{C}_j$, $(\jmath_1, \ldots, \jmath_k) \in \mathscr{C}_k$, and $(\jmath_{k+1}, \ldots, \jmath_i) \in \mathscr{P}_{i-k}$. This implies that $(\jmath_{k+1}, \ldots, \jmath_j) \in \mathscr{P}_{j-k}$ because the first $j-k$ terms of a sequence of length $i-k > j-k$ in $\mathscr{P}_{i-k}$ must necessarily belong to $\mathscr{P}_{j-k}$ given the definition in [A.1.4]. Thus it follows that there exists a $(\jmath_1, \ldots, \jmath_j) \in \mathscr{C}_j \cap (\mathscr{C}_k \times \mathscr{P}_{j-k})$ for some $k < j$. However, this directly contradicts the definition of $\mathscr{C}_j$ in [A.1.7]. Therefore, [A.1.8] must be true.

Given the recursion for $\{\phi_i\}_{i=1}^{\infty}$ in [A.1.5] and the expression for ψ_i in [A.1.6], the following provides a formula for ϕ_i:

$$\phi_i = \left\{(-1)^i \sum_{(\jmath_1,\ldots,\jmath_i)\in\mathscr{P}_i} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell}\right\} - \sum_{j=1}^{i-1} \phi_j \left\{(-1)^{i-j} \sum_{(\jmath_1,\ldots,\jmath_{i-j})\in\mathscr{P}_{i-j}} \prod_{\ell=1}^{i-j} \sigma_{\jmath_\ell}\right\}. \qquad \text{[A.1.9]}$$

It is claimed that the following equation holds for all $i = 1, 2, \ldots$:

$$\phi_i = (-1)^i \sum_{(\jmath_1,\ldots,\jmath_i)\in\mathscr{C}_i} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell}, \tag{A.1.10}$$

Suppose this statement has already been proved for $j = 1, \ldots, i-1$ and substitute it into [A.1.9] to obtain:

$$\phi_i = \left((-1)^i \sum_{(\jmath_1,\ldots,\jmath_i)\in\mathscr{P}_i} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell} \right) - \sum_{j=1}^{i-1} \left((-1)^i \sum_{(\jmath_1,\ldots,\jmath_i)\in(\mathscr{C}_j\times\mathscr{P}_{i-j})} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell} \right), \tag{A.1.11}$$

where the following has been used:

$$\left\{ (-1)^j \sum_{(\jmath_1,\ldots,\jmath_j)\in\mathscr{C}_j} \prod_{\ell=1}^{j} \sigma_{\jmath_\ell} \right\} \left\{ (-1)^{i-j} \sum_{(\jmath_1,\ldots,\jmath_{i-j})\in\mathscr{P}_{i-j}} \prod_{\ell=1}^{i-j} \sigma_{\jmath_\ell} \right\} = (-1)^i \sum_{(\jmath_1,\ldots,\jmath_i)\in(\mathscr{C}_j\times\mathscr{P}_{i-j})} \prod_{\ell=1}^{i} \sigma_{\jmath_\ell}.$$

It follows from [A.1.11] that [A.1.10] holds for i if the sets $\mathscr{C}_j \times \mathscr{P}_{i-j}$ and $\mathscr{C}_k \times \mathscr{P}_{i-k}$ are disjoint for all $j \neq k$, which is the claim [A.1.8] established earlier. Now note that the definitions of $\mathscr{C}_1$, σ_1 and ϕ_1 imply that [A.1.10] holds for $i = 1$. Therefore, the expression for ϕ_i in [A.1.10] is verified for all i by induction. Since $\varphi_i = (-1)\phi_{i+1}$ by definition, equation [A.1.3] is demonstrated for the particular sets $\{\mathscr{C}_i\}_{i=2}^{\infty}$ characterized in [A.1.7]. This completes the proof. ∎

A.2 Proof of Proposition 1

Let $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ denote the z-transform of the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$. Lemma 1 demonstrates that $\psi(z)$ is analytic on $\mathscr{D}_\rho \equiv \{ z \in \mathbb{C} \mid |z| \leq \rho \}$ for some $\rho > 1$. Since $1 \in \mathscr{D}_\rho$, it follows that $\psi(1) = \sum_{\ell=0}^{\infty} \psi_\ell$ is finite (and positive given that $\psi_0 = 1$ and $\psi_\ell \geq 0$). Define $\omega_0 = \psi(1)^{-1}$ and $\omega_\ell = \omega_0 \psi_\ell$ for $\ell \geq 1$. By construction, the sequence $\{\omega_\ell\}_{\ell=0}^{\infty}$ satisfies $\sum_{\ell=0}^{\infty} \omega_\ell = 1$, and $\omega_\ell = (1-\alpha_\ell)\omega_{\ell-1}$ since $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$. Note also that

$$\sum_{\ell=1}^{\infty} \alpha_\ell \omega_{\ell-1} = \omega_0 \sum_{\ell=1}^{\infty} \alpha_\ell \psi_{\ell-1} = \omega_0 \sum_{\ell=1}^{\infty} (\psi_{\ell-1} - \psi_\ell) = \omega_0,$$

as $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$ and $\psi_0 = 1$. This confirms that $\{\omega_\ell\}_{\ell=0}^{\infty}$ is a stationary age distribution. There can be only one such distribution because $\{\omega_\ell\}_{\ell=0}^{\infty}$ must satisfy $\omega_\ell = (1-\alpha_\ell)\omega_{\ell-1}$ for all $\ell \geq 1$. This leaves only ω_0 to be determined, but this is pinned down by the requirement $\sum_{\ell=0}^{\infty} \omega_\ell = 1$.

Now suppose that $\alpha_\ell \geq \underline{\alpha}$ for all ℓ for some $\underline{\alpha}$ satisfying $0 < \underline{\alpha} < 1$. Since $\psi_{\ell+1} = (1-\alpha_{\ell+1})\psi_\ell$, this implies $0 \leq \psi_{\ell+1} \leq (1-\underline{\alpha})\psi_\ell$ for all ℓ. Hence Lemma 2 implies that there exists a $\rho > 1$ such that $\psi(z) = 0$ has no roots on $\mathscr{D}_\rho$. Lemma 1 shows that this condition implies that the stationary age distribution is stable, completing the proof.

A.3 Proof of Proposition 2

The first step is to derive the standard representation of the Phillips curve [2.6] from equations [2.3], [2.4] and [2.5]. Let $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ and $\omega(z) \equiv \sum_{\ell=0}^{\infty} \omega_\ell z^\ell$ be the z-transforms of the sequences of survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ and the age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$. Written in terms of the lag and forward operators $\mathbb{L}$ and $\mathbb{F}$, equations [2.4] and [2.5] become:

$$\mathsf{r}_t = \psi(\beta)^{-1} \mathbb{E}_t \left[\psi(\beta\mathbb{F}) \mathsf{p}_t^* \right], \quad \text{and} \quad \mathsf{p}_t = \omega(\mathbb{L}) \mathsf{r}_t. \tag{A.3.1}$$

Note that $\omega_\ell = \psi_\ell \omega_0$, so $\omega_0 = \psi(1)^{-1}$ since $\omega(1) = 1$. This justifies the relationship $\omega(z) = \psi(1)^{-1}\psi(z)$ between $\omega(z)$ and $\psi(z)$. By using this result, eliminating the reset price r_t from [A.3.1], and substituting the expression for p_t^* from [2.3]:

$$\left\{\mathbb{I} - \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{p}_t = \nu\left\{\psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{x}_t, \qquad \text{[A.3.2]}$$

where $\mathbb{I}$ denotes the identity operator.

The left-hand side of [A.3.2] is

$$\left\{\mathbb{I} - \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{p}_t = \mathsf{p}_t - \frac{\sum_{j=0}^{\infty}\psi_j\sum_{\ell=0}^{\infty}\beta^\ell\psi_\ell\mathbb{E}_{t-j}\mathsf{p}_{t-j+\ell}}{\sum_{j=0}^{\infty}\psi_j\sum_{\ell=0}^{\infty}\beta^\ell\psi_\ell}. \qquad \text{[A.3.3]}$$

The definition of inflation $\pi_t = \mathsf{p}_t - \mathsf{p}_{t-1}$ implies $\mathsf{p}_{t-j+\ell} = \mathsf{p}_{t-j} + \pi_{t-j+1} + \cdots + \pi_{t-j+\ell}$, so

$$\left\{\mathbb{I} - \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{p}_t = \frac{\sum_{\ell=0}^{\infty}\psi_\ell(\mathsf{p}_t - \mathsf{p}_{t-\ell})}{\sum_{\ell=0}^{\infty}\psi_\ell} - \frac{\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_j\left(\sum_{i=\ell}^{\infty}\beta^i\psi_i\right)\mathbb{E}_{t-j}\pi_{t-j+\ell}}{\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\beta^\ell\psi_j\psi_\ell}.$$

The definition of inflation also implies $\mathsf{p}_t - \mathsf{p}_{t-\ell} = \pi_{t-\ell+1} + \cdots + \pi_t$, thus

$$\begin{aligned}\left\{\mathbb{I} - \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{p}_t = {}& \frac{\sum_{\ell=1}^{\infty}\psi_\ell}{\sum_{\ell=0}^{\infty}\psi_\ell}\pi_t + \frac{\sum_{\ell=1}^{\infty}\left(\sum_{i=\ell+1}^{\infty}\psi_i\right)\pi_{t-\ell}}{\sum_{\ell=0}^{\infty}\psi_\ell} \\ & - \frac{\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_j\left(\sum_{i=\ell}^{\infty}\beta^i\psi_i\right)\mathbb{E}_{t-j}\pi_{t-j+\ell}}{\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\beta^\ell\psi_j\psi_\ell}. \qquad \text{[A.3.4]}\end{aligned}$$

The right-hand side of [A.3.2] is

$$\nu\left\{\psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{x}_t = \nu\frac{\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\beta^\ell\psi_j\psi_\ell\mathbb{E}_{t-j}\mathsf{x}_{t-j+\ell}}{\sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\beta^\ell\psi_j\psi_\ell}. \qquad \text{[A.3.5]}$$

Using the expressions in [A.3.4] and [A.3.5], it is seen that [A.3.2] is equivalent to the standard Phillips curve equation [2.6] with the coefficients:

$$\mathrm{a}_\ell = -\frac{\sum_{i=\ell+1}\psi_i}{\sum_{i=1}^{\infty}\psi_i}, \quad \mathrm{b}_{j\ell} = \frac{\psi_j\sum_{i=\ell}^{\infty}\psi_i}{\sum_{i=1}^{\infty}\sum_{h=0}^{\infty}\beta^h\psi_i\psi_h}, \quad \text{and} \quad \mathrm{c}_{j\ell} = \frac{\beta^\ell\psi_j\psi_\ell}{\sum_{i=1}^{\infty}\sum_{h=0}^{\infty}\beta^h\psi_i\psi_h}.$$

Now suppose the hazard function implies that the stationary age distribution of prices is stable. As Lemma 1 shows, this is equivalent to there being a $\rho > 1$ such that $\psi(z)$ has no roots in the set $\mathscr{D}_\rho \equiv \left\{z \in \mathbb{C} \mid |z| \leq \rho\right\}$. Under this condition, the function $\phi(z) \equiv \psi(z)^{-1}$ is analytic on $\mathscr{D}_\rho$, which is equivalent to $\phi(z)$ being equal to its Taylor expansion around $z = 0$ for all $z \in \mathscr{D}_\rho$. Thus, $\phi(z) \equiv 1 - \sum_{\ell=1}^{\infty}\phi_\ell z^\ell$ for some sequence of numbers $\{\phi_\ell\}_{\ell=1}^{\infty}$, with $\sum_{\ell=1}^{\infty}|\phi_\ell| < \infty$ since $\mathscr{D}_\rho$ encloses the unit circle. The first term in the Taylor series of $\phi(z)$ is 1 because $\psi(0) = \psi_0 = 1$.

Since $\phi(z)\psi(z) = 1$ for all $|z| \leq 1$, it follows that $\mathbb{I} = \psi(\mathbb{L})\phi(\mathbb{L})$, which allows the left-hand side of [A.3.2] to be expressed equivalently as follows:

$$\left\{\mathbb{I} - \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{p}_t = \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\left\{\psi(1)\psi(\beta)\phi(\mathbb{L}) - \mathbb{E}_t\psi(\beta\mathbb{F})\right\}\mathsf{p}_t. \qquad \text{[A.3.6]}$$

It also follows from $\phi(z)\psi(z) = 1$ that $\mathbb{I} = \psi(\beta\mathbb{F})\phi(\beta\mathbb{F})$, and thus $\phi(\mathbb{L}) = \mathbb{I}\phi(\mathbb{L}) = \psi(\beta\mathbb{F})\phi(\beta\mathbb{F})\phi(\mathbb{L})$. Furthermore, note that the power series $\phi(\mathbb{L}) \equiv \sum_{\ell=0}^{\infty}\phi_\ell\mathbb{L}^\ell$ contains only non-negative powers of the lag operator $\mathbb{L}$, so $\phi(\mathbb{L})\mathsf{p}_t = \mathbb{E}_t\phi(\mathbb{L})\mathsf{p}_t$. Putting these two results together implies $\phi(\mathbb{L})\mathsf{p}_t = \mathbb{E}_t\psi(\beta\mathbb{F})\phi(\beta\mathbb{F})\phi(\mathbb{L})\mathsf{p}_t$. Then observe that because the power series $\psi(\beta\mathbb{F}) \equiv \sum_{\ell=0}^{\infty}\beta^\ell\psi_\ell\mathbb{F}^\ell$ contains only non-negative powers of $\mathbb{F}$, the law of iterated expectations (from which it follows that the conditional expectation operator $\mathbb{E}_t$ commutes with all non-negative powers of the forward operator $\mathbb{F}$) implies

$$\phi(\mathbb{L})\mathsf{p}_t = \mathbb{E}_t\left[\psi(\beta\mathbb{F})\left\{\mathbb{E}_t\phi(\beta\mathbb{F})\phi(\mathbb{L})\right\}\mathsf{p}_t\right].$$

This result, together with [A.3.6], and noting $\phi(\beta\mathbb{F})\phi(\mathbb{L}) = \phi(\mathbb{L})\phi(\beta\mathbb{F})$, $\psi(1) = \phi(1)^{-1}$ and $\psi(\beta) = \phi(\beta)^{-1}$, yields

$$\left\{\mathbb{I} - \psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\} \mathsf{p}_t = \left\{\psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\} \mathbb{E}_t \left\{\phi(1)^{-1}\phi(\beta)^{-1}\phi(\mathbb{L})\phi(\beta\mathbb{F}) - \mathbb{I}\right\} \mathsf{p}_t. \quad \text{[A.3.7]}$$

Equating this expression to the right-hand side of [A.3.2] leads to the following equation that is exactly equivalent to the Phillips curve [2.6]:

$$\left\{\psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\} \left(\mathbb{E}_t \left[\left\{\phi(1)^{-1}\phi(\beta)^{-1}\phi(\mathbb{L})\phi(\beta\mathbb{F}) - \mathbb{I}\right\} \mathsf{p}_t\right] - \nu \mathsf{x}_t\right) = 0. \quad \text{[A.3.8]}$$

Now define the function $\chi(z) \equiv \phi(1)^{-1}\phi(\beta)^{-1}\phi(z)\phi(\beta z^{-1}) - 1$, which is analytic on the annulus $\mathscr{A}_\rho \equiv \left\{z \in \mathbb{C} \mid \beta\rho^{-1} \leq |z| \leq \rho\right\}$ given that $\phi(z)$ is analytic and has no roots on $\mathscr{D}_\rho$. Notice that $\chi(1) = 0$, so it follows that there is another function $\theta(z)$ analytic on $\mathscr{A}_\rho$ such that $\chi(z) = (1 - z)\theta(z)$. The function $\theta(z)$ is equal to its Laurent series expansion $\theta(z) = \sum_{\ell \to -\infty}^{\infty} \theta_\ell z^\ell$ for all $z \in \mathscr{A}_\rho$. Since $\mathscr{A}_\rho$ includes the unit circle, it follows that $\sum_{\ell \to -\infty}^{\infty} |\theta_\ell| < \infty$. Make the following definitions of sequences $\{\lambda_\ell\}_{\ell=1}^{\infty}$ and $\{\xi_\ell\}_{\ell=1}^{\infty}$, and coefficient κ appearing in the new Phillips curve [2.7]:

$$\lambda_\ell \equiv -\frac{\theta_\ell}{\theta_0}, \quad \xi_\ell \equiv -\frac{\theta_{-\ell}}{\theta_0}, \quad \text{and} \quad \kappa \equiv \frac{1}{\theta_0}.$$

With these definitions, the sequences clearly satisfy $\sum_{\ell=1}^{\infty} |\lambda_\ell| < \infty$ and $\sum_{\ell=1}^{\infty} |\xi_\ell| < \infty$ (it can be shown that $\theta_0 \neq 0$ using the argument presented in the proof of Proposition 6).

Now define

$$\mathsf{d}_t \equiv \pi_t - \sum_{\ell=1}^{\infty} \lambda_\ell \pi_{t-\ell} - \sum_{\ell=1}^{\infty} \xi_\ell \mathbb{E}_t \pi_{t+\ell} - \nu\kappa \mathsf{x}_t, \quad \text{[A.3.9]}$$

and note that the definitions above imply $\mathsf{d}_t = \kappa\left\{\mathbb{E}_t\left[\theta(\mathbb{L})\pi_t\right] - \nu \mathsf{x}_t\right\}$. Since $\pi_t = (\mathbb{I} - \mathbb{L})\mathsf{p}_t$ and $\chi(\mathbb{L}) = (\mathbb{I} - \mathbb{L})\theta(\mathbb{L})$, it follows that $\theta(\mathbb{L})\pi_t = \chi(\mathbb{L})\mathsf{p}_t$ and hence $\mathsf{d}_t = \kappa\left\{\mathbb{E}_t\left[\chi(\mathbb{L})\mathsf{p}_t\right] - \nu \mathsf{x}_t\right\}$. Therefore, comparing this expression for d_t to equation [A.3.8], the Phillips curve [2.6] is equivalent to $\left\{\psi(1)^{-1}\psi(\beta)^{-1}\psi(\mathbb{L})\mathbb{E}_t\psi(\beta\mathbb{F})\right\} \mathsf{d}_t = 0$, and thus to

$$\psi(\mathbb{L})\mathbb{E}_t\left[\psi(\beta\mathbb{F})\mathsf{d}_t\right] = 0, \quad \text{[A.3.10]}$$

holding in all time periods t. Let $\mathsf{e}_t \equiv \mathbb{E}_t\left[\psi(\beta\mathbb{F})\mathsf{d}_t\right]$, with equation [A.3.10] being equivalent to $\psi(\mathbb{L})\mathsf{e}_t = 0$ for all t.

Note that by comparing [A.3.9] to [2.7], the new Phillips curve equation is equivalent to $\mathsf{d}_t = 0$ for all t. Suppose the new Phillips curve equation [2.7] holds. Thus $\mathsf{d}_t = 0$ for all t and hence $\mathsf{e}_t = 0$ for all t as well. It follows that $\psi(\mathbb{L})\mathsf{e}_t = 0$, so the original Phillips curve [2.6] must hold.

Conversely, suppose the original Phillips curve [2.6] holds, which implies $\psi(\mathbb{L})\mathsf{e}_t = 0$ using [A.3.10]. Given the stability of the stationary age distribution, it has been shown that $\psi(z) = 0$ has no roots on or inside the unit circle. Thus if $\mathsf{e}_{t_0} \neq 0$ for some t_0, it follows from $\psi(\mathbb{L})\mathsf{e}_t = 0$ that e_t is unbounded for time periods before t_0. Now given the location of the roots of $\psi(z) = 0$, it follows from $0 < \beta < 1$ that $\psi(\beta z) = 0$ has no roots on or inside the unit circle. Hence if $\mathsf{e}_t = 0$ for all t, the only bounded solution of $\mathsf{e}_t \equiv \mathbb{E}_t\left[\psi(\beta\mathbb{F})\mathsf{d}_t\right]$ is $\mathsf{d}_t = 0$ for all t. On the other hand, if e_t is unbounded over all time periods t, then d_t must also be unbounded. If d_t is unbounded then equation [A.3.9] shows that either inflation π_t or real marginal cost x_t must be unbounded. Consequently, if attention is restricted to bounded rational expectations solutions (as is conventional), the original Phillips curve [2.6] implies $\mathsf{e}_t = 0$ for all t, and hence $\mathsf{d}_t = 0$ for all t. This then demonstrates that the new Phillips curve [2.7] must hold, completing the proof.

A.4 Proof of Proposition 3

Let $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ denote the z-transform of the sequence of survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ generated by some hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ from parameters α and $\{\varphi_i\}_{i=1}^{n}$ using the recursion [3.1]. Define the polynomial

$$\phi(z) = 1 - \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i\right) z + \sum_{\jmath=1}^{n} \varphi_\jmath z^{\jmath+1} \qquad \text{[A.4.1]}$$

using these parameters. Since the recursion in [3.1] is equivalent to [3.2], by multiplying the power series $\phi(z)$ and $\psi(z)$ and noting that $\psi_0 = 1$, it follows that $\phi(z)\psi(z) = 1$ for all z for which $\psi(z)$ is analytic.

The hazard function implies a unique stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$, with its z-transform denoted by $\omega(z) \equiv \sum_{\ell=0}^{\infty} \omega_\ell z^\ell$. Since $\omega_\ell = (1-\alpha_\ell)\omega_{\ell-1}$ and $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$, it follows that $\omega(z)$ is a multiple of $\psi(z)$. In particular, as $\psi_0 = 1$, it must be the case that $\omega(z) = \omega_0\psi(z)$. As $\{\omega_\ell\}_{\ell=0}^{\infty}$ is a probability distribution, it follows that $\omega(1) = 1$, and thus $\omega_0 = \psi(1)^{-1}$ and $\omega(z) = \psi(1)^{-1}\psi(z)$. Together with $\phi(z)\psi(z) = 1$, it is established that $\phi(z)\omega(z) = \psi(1)^{-1}$. Since $\psi(1)^{-1} = \phi(1)$:

$$\omega(z) = \phi(1)\phi(z)^{-1}. \qquad \text{[A.4.2]}$$

(i) Let $\bar{\alpha}$ denote the average probability of price adjustment, calculated with respect to the stationary age distribution of prices at the beginning of any period. This distribution is given by $\{\omega_{\ell-1}\}_{\ell=1}^{\infty}$, so $\bar{\alpha} = \sum_{\ell=1}^{\infty} \omega_{\ell-1}\alpha_\ell$. Using the fact that $\omega_\ell = (1-\alpha_\ell)\omega_{\ell-1}$, it follows that $\omega_{\ell-1}\alpha_\ell = \omega_{\ell-1} - \omega_\ell$ and thus

$$\bar{\alpha} = \sum_{\ell=1}^{\infty}(\omega_{\ell-1} - \omega_\ell) = \sum_{\ell=0}^{\infty} \omega_\ell - \sum_{\ell=1}^{\infty} \omega_\ell = \omega_0. \qquad \text{[A.4.3]}$$

The fraction of newly set prices is ω_0. Since $\omega_0 = \omega(0)$ and $\phi(0) = 1$, it follows from [A.4.1] and [A.4.2] that

$$\bar{\alpha} = \omega_0 = \phi(1) = \alpha, \qquad \text{[A.4.4]}$$

for all values of $\{\varphi_i\}_{i=1}^{n}$.

(ii) Now consider the expected duration of a newly set price. If $\varsigma_\ell \equiv 1 - \alpha_\ell$ denotes the probability of price stickiness in the current period if ℓ periods have elapsed since the last change then $\alpha_\ell \prod_{\jmath=1}^{\ell-1} \varsigma_\jmath$ is the probability that a price will survive for exactly ℓ periods after first being set before being changed. The expected duration is denoted by $\hbar$:

$$\hbar \equiv \sum_{\ell=1}^{\infty} \ell\alpha_\ell \prod_{\jmath=1}^{\ell-1} \varsigma_\jmath.$$

The definition of the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ implies $\psi_{\ell-1} = \prod_{\jmath=1}^{\ell-1} \varsigma_\jmath$. Together with $\alpha_\ell\psi_{\ell-1} = \psi_{\ell-1} - \psi_\ell$, the expected duration is given by

$$\hbar = \sum_{\ell=1}^{\infty} \ell\alpha_\ell\psi_{\ell-1} = \sum_{\ell=1}^{\infty} \ell(\psi_{\ell-1} - \psi_\ell) = \sum_{\ell=0}^{\infty}(\ell+1)\psi_\ell - \sum_{\ell=0}^{\infty} \ell\psi_\ell = \sum_{\ell=0}^{\infty} \psi_\ell = \psi(1). \qquad \text{[A.4.5]}$$

As $\phi(z)\psi(z) = 1$, it follows that $\psi(1) = \phi(1)^{-1}$. The result in [A.4.4] then implies that $\hbar = \alpha^{-1}$.

(iii) Let $\hbar_\alpha$ denote the average age of the prices that are changed. Using Bayes' law, the probability that a price has age ℓ conditional on being changed is the product of α_ℓ and $\omega_{\ell-1}$ divided by $\alpha = \omega_0$. Since $\omega_{\ell-1}/\omega_0 = \psi_{\ell-1}$, it follows that $\hbar_\alpha$ is given by

$$\hbar_\alpha = \sum_{\ell=1}^{\infty} \ell\alpha_\ell\psi_{\ell-1}. \qquad \text{[A.4.6]}$$

The result in [A.4.5] then implies $\hbar_\alpha = \hbar = \alpha^{-1}$.

Let $\hbar_\varsigma$ denote the average age of the prices that are not changed. Again, using Bayes' law, the probability that a price has age ℓ conditional on not being changed is the product of $\varsigma_\ell = 1 - \alpha_\ell$ and $\omega_{\ell-1}$ divided by

$1-\alpha$. Thus $\hbar_\varsigma$ is given by

$$\hbar_\varsigma = \sum_{\ell=1}^{\infty} \ell \frac{\varsigma_\ell \omega_{\ell-1}}{1-\alpha}.$$

Using $\varsigma_\ell = 1-\alpha_\ell$ and $\omega_{\ell-1} = \alpha\psi_{\ell-1}$ since $\omega_0 = \alpha$:

$$\hbar_\varsigma = \frac{1}{1-\alpha}\left(\sum_{\ell=1}^{\infty} \ell\omega_{\ell-1} - \alpha\sum_{\ell=1}^{\infty} \ell\alpha_\ell\psi_{\ell-1}\right) = \frac{1}{1-\alpha}\left(\sum_{\ell=0}^{\infty} \omega_\ell + \sum_{\ell=0}^{\infty} \ell\omega_\ell - \alpha\sum_{\ell=1}^{\infty} \ell\alpha_\ell\psi_{\ell-1}\right).$$

Note that $\sum_{\ell=0}^{\infty} \omega_\ell = 1$. From the definition of $\omega(z)$ it follows that $\omega'(z) = \sum_{\ell=0}^{\infty} \ell\omega_\ell z^{\ell-1}$ and thus $\omega'(1) = \sum_{\ell=0}^{\infty} \ell\omega_\ell$. Substituting these results and using the expression for $\hbar_\alpha$ from [A.4.6] together with $\hbar_\alpha = \alpha^{-1}$ to deduce:

$$\hbar_\varsigma = \frac{1}{1-\alpha}\left(1+\omega'(1) - \alpha\alpha^{-1}\right) = \frac{\omega'(1)}{1-\alpha}. \qquad \text{[A.4.7]}$$

Differentiation of both sides of [A.4.2] yields:

$$\omega'(z) = -\frac{\phi(1)\phi'(z)}{\phi(z)^2},$$

and hence $\omega'(1) = -\phi'(1)\phi(1)^{-1}$. Differentiation of the polynomial $\phi(z)$ in [A.4.1] implies $\phi'(z) = -\left(1-\alpha+\sum_{i=1}^{n}\varphi_i\right) + \sum_{j=1}^{n}(j+1)\varphi_j z^j$, from which it follows that $\phi'(1) = -\left(1-\alpha-\sum_{i=1}^{n} i\varphi_i\right)$. And since $\phi(1) = \alpha$:

$$\omega'(1) = \left(1-\alpha-\sum_{i=1}^{n} i\varphi_i\right)\alpha^{-1}. \qquad \text{[A.4.8]}$$

Therefore, using [A.4.7], the difference between the average ages of prices conditional on adjustment and non-adjustment is

$$\hbar_\alpha - \hbar_\varsigma = \alpha^{-1} - \left(1-\alpha-\sum_{i=1}^{n} i\varphi_i\right)\alpha^{-1}(1-\alpha)^{-1} = \left(\sum_{i=1}^{n} i\varphi_i\right)\alpha^{-1}(1-\alpha)^{-1}.$$

(iv) Let $\hbar \equiv \sum_{\ell=0}^{\infty} \ell\omega_\ell$ denote the average age of prices actually in use according to the stationary distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$. Using the definition of $\omega(z)$ it follows that $\hbar = \omega'(1)$. Hence, [A.4.8] implies

$$\hbar = \left(1-\alpha-\sum_{i=1}^{n} i\varphi_i\right)\alpha^{-1} = \left(1-\sum_{i=1}^{n} i\varphi_i\right)\alpha^{-1} - 1. \qquad \text{[A.4.9]}$$

(v) The hazard function recursion [3.1] implies that the probability of adjusting the most recently set price is

$$\alpha_1 = \alpha - \sum_{i=1}^{n}\varphi_i.$$

So α_1 is clearly strictly decreasing in each φ_i.

Let $\alpha_\infty \equiv \lim_{\ell\to\infty}\alpha_\ell$ be the limiting value of the hazard function for price spells of arbitrarily long duration. The recursion for the hazard function is equivalent to the linear recursion for the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ in [3.2]. The recursion [3.2] is a linear difference equation with $\phi(z^{-1}) = 0$ in [A.4.1] being the characteristic polynomial (since $\phi(z)\psi(z) = 1$).

Now consider parameter values α and $\{\varphi_i\}_{i=1}^{n}$ such that $\phi(z) = 0$ has no repeated roots. This will be without loss of generality because there is always a set of parameters implying no repeated roots arbitrarily close to parameters for which there are repeated roots. With no repeated roots, the solution for the sequence of survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ takes the following general form

$$\psi_\ell = \sum_{\jmath=1}^{n+1} \varkappa_\jmath \zeta_\jmath^\ell, \tag{A.4.10}$$

for some sequence of coefficients $\{\varkappa_\jmath\}_{\jmath=1}^{n+1}$, and a sequence $\{\zeta_\jmath\}_{\jmath=1}^{n+1}$ where each $\zeta_\jmath$ is a reciprocal of one of the $n+1$ distinct roots of $\phi(z)=0$, that is, $\phi(\zeta_\jmath^{-1})=0$.

Without loss of generality, order the sequence $\{\zeta_\jmath\}_{\jmath=1}^{n+1}$ so that $|\zeta_1| \geq |\zeta_2| \geq \cdots \geq |\zeta_{n+1}|$. As $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$, it follows that $\alpha_\ell = 1-(\psi_\ell/\psi_{\ell-1})$ and hence:

$$\alpha_\ell = 1 - \frac{\sum_{\jmath=1}^{n+1} \varkappa_\jmath \zeta_\jmath^\ell}{\sum_{\jmath=1}^{n+1} \varkappa_\jmath \zeta_\jmath^{\ell-1}} = 1 - \frac{\zeta_1 + \sum_{\jmath=2}^{n+1} \zeta_\jmath \frac{\varkappa_\jmath}{\varkappa_1}\left(\frac{\zeta_\jmath}{\zeta_1}\right)^{\ell-1}}{1+\sum_{\jmath=2}^{n+1} \frac{\varkappa_\jmath}{\varkappa_1}\left(\frac{\zeta_\jmath}{\zeta_1}\right)^{\ell-1}}.$$

With no repeated roots, $\zeta_1 \neq \zeta_2$, so a necessary condition for the limit $\lim_{\ell\to\infty}\alpha_\ell$ to exist is that $|\zeta_1| > |\zeta_2|$ (using the ordering of the roots), which also requires ζ_1 to be a real number. Under this condition, $\alpha_\infty \equiv \lim_{\ell\to\infty}\alpha_\ell = 1-\zeta_1$. For this limit to be economically meaningful and ensure $\alpha_\infty > 0$, it is necessary that $0 \leq \zeta_1 < 1$.

It is known that $\phi(z)\psi(z)=1$, so $\phi(1)=\psi(1)^{-1}$, which is necessarily positive since $\psi_0=1$ and $\psi_\ell \geq 0$ for all ℓ. As ζ_1 is the largest of the reciprocals of the roots of $\phi(z)=0$, there must be no value of ζ between ζ_1 and 1 such that $\phi(\zeta^{-1})=0$. Since $\phi(z)$ in [A.4.1] is a polynomial, it is a continuous function. Together with $\phi(1)>0$ and the absence of any value of ζ between ζ_1 and 1 such that $\phi(\zeta^{-1})=0$, it must be the case that $\phi'(\zeta_1^{-1})<0$.

The value of ζ_1 is characterized by $\phi(\zeta_1^{-1})=0$, so the change in ζ_1 resulting from a change in a parameter φ_i is implicitly determined by the condition $\phi(\zeta_1^{-1})=0$. Differentiating this condition yields

$$\left.\frac{\partial\zeta_1^{-1}}{\partial\varphi_i}\right|_{\phi(\zeta_1^{-1})=0} = -\frac{1}{\zeta_1^{i+1}\phi'(\zeta_1^{-1})}. \tag{A.4.11}$$

As $\alpha_\infty = 1-(\zeta_1^{-1})^{-1}$, it follows that $\partial\alpha_\infty/\partial\zeta_1^{-1} = \zeta_1^2$, and thus using the chain rule with [A.4.11]:

$$\frac{\partial\alpha_\infty}{\partial\varphi_i} = -\frac{1}{\zeta_1^{i-1}\phi'(\zeta_1^{-1})} > 0,$$

since $\phi'(\zeta_1^{-1})<0$ as demonstrated above.

(vi) In what follows, suppose that $n=\infty$ in the hazard function recursion [3.1]. This is without loss of generality because any superfluous φ_i parameters can be set to zero. Equation [3.1] implies

$$\alpha_{\ell+1}-\alpha_\ell = \sum_{i=1}^{\ell}\varphi_i\left(\prod_{j=\ell+1-i}^{\ell}(1-\alpha_j)\right)^{-1} - \sum_{i=1}^{\ell-1}\varphi_i\left(\prod_{j=\ell-i}^{\ell-1}(1-\alpha_j)\right)^{-1},$$

and by combining overlapping terms and extracting common factors:

$$\alpha_{\ell+1}-\alpha_\ell = \varphi_\ell\left(\prod_{j=1}^{\ell}(1-\alpha_j)\right)^{-1} + \sum_{i=1}^{\ell-1}\varphi_i\left(\prod_{j=\ell-i}^{\ell}(1-\alpha_j)\right)^{-1}\{(1-\alpha_{\ell-i})-(1-\alpha_\ell)\}.$$

Therefore, the change in the hazard function is given by:

$$\alpha_{\ell+1}-\alpha_\ell = \sum_{i=1}^{\ell-1}\varphi_i(\alpha_\ell-\alpha_{\ell-i})\left(\prod_{j=\ell-i}^{\ell}(1-\alpha_j)\right)^{-1} + \varphi_\ell\left(\prod_{j=1}^{\ell}(1-\alpha_j)\right)^{-1}. \tag{A.4.12}$$

It follows that $\varphi_i = 0$ for all i implies $\alpha_\ell = \alpha$ for all ℓ. Similarly, suppose $\alpha_\ell = \alpha_1$ for all ℓ. It follows from [A.4.12] that $\varphi_i = 0$ for all i.

(vii) Suppose that $\varphi_i \geq 0$ for all i. It follows immediately from [A.4.12] that $\alpha_2 \geq \alpha_1$. Now suppose that $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{i-1} \leq \alpha_\ell$ has already been established for some ℓ. Given this supposition, it follows that $\alpha_\ell - \alpha_{\ell-i} \geq 0$ for all $i = 1, \ldots, \ell - 1$. Equation [A.4.12] then implies that $\alpha_{\ell+1} \geq \alpha_\ell$. This proves $\alpha_{\ell+1} \geq \alpha_\ell$ for all ℓ by induction.

(viii) Define the sequence $\{\sigma_\ell\}_{\ell=1}^{\infty}$ using $\sigma_1 = -(1-\alpha_1)$ and $\sigma_\ell = \alpha_\ell - \alpha_{\ell-1}$ for $\ell \geq 2$. If $\alpha_{\ell+1} \leq \alpha_\ell$ for all ℓ then $\sigma_\ell \leq 0$ for all ℓ. It follows from the expression for φ_i in equation [A.1.3] justified by Lemma 3 that φ_i is the product of $(-1)^i$ and $i+1$ non-positive terms. Hence, $\varphi_i \leq 0$ for all i is established. This completes the proof.

A.5 Proof of Proposition 4

Let $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ denote the z-transform of the sequence of survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ generated by a hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$. If the hazard function implies the stationary age distribution is stable then Lemma 1 shows there exists a $\rho > 1$ such that $\psi(z)$ has no roots in $\mathscr{D}_\rho \equiv \left\{z \in \mathbb{C} \,\middle|\, |z| \leq \rho\right\}$. Define the function $\phi(z) \equiv \psi(z)^{-1}$ on $\mathscr{D}_\rho$, which is analytic because $\psi(z) \neq 0$ for all $z \in \mathscr{D}_\rho$.

Since $\phi(z)$ is an analytic function, it is equal to its Taylor series expansion around $z = 0$ (contained in $\mathscr{D}_\rho$). Thus $\phi(z) \equiv 1 - \sum_{i=1}^{\infty} \phi_i z^\ell$ for some sequence $\{\phi_i\}_{i=1}^{\infty}$ (the leading term of the Taylor series is 1 because $\psi(0) = \psi_0 = 1$). As $z = 1$ belongs to $\mathscr{D}_\rho$, it follows that $\sum_{i=1}^{\infty} |\phi_i| < \infty$.

The definition of $\phi(z)$ requires $\phi(z)\psi(z) = 1$ for all $z \in \mathscr{D}_\rho$. Multiplying the power series for $\phi(z)$ and $\psi(z)$ yields

$$\phi(z)\psi(z) = \psi_0 + \sum_{\ell=1}^{\infty} \left(\psi_\ell - \sum_{i=1}^{\ell} \phi_i \psi_{\ell-i} \right) z^\ell.$$

Since $\psi_0 = 1$ always, $\phi(z)\psi(z) = 1$ holds for all $z \in \mathscr{D}_\rho$ if and only if $\psi_\ell = \sum_{i=1}^{\ell} \phi_i \psi_{\ell-i}$ is true for all ℓ. Define α and $\{\varphi_i\}_{i=1}^{\infty}$ according to $\alpha \equiv 1 - \sum_{i=1}^{\infty} \phi_i$ and $\varphi_i \equiv -\phi_{i+1}$. With these definitions, the recursion for $\{\psi_\ell\}_{\ell=0}^{\infty}$ in [3.2] holds with $n = \infty$, which is equivalent to the original recursion for the hazard function in [3.1]. Given the definitions, it has also been shown that $\sum_{i=1}^{\infty} |\varphi_i| < \infty$. This completes the proof.

A.6 Proof of Proposition 5

(i) Define the sequence of probabilities of price stickiness $\{\varsigma_\ell\}_{\ell=1}^{\infty}$ as $\varsigma_\ell \equiv 1 - \alpha_\ell$ using the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$. If the parameters α and $\{\varphi_i\}_{i=1}^{n}$ generate a well-defined hazard function then it follows that $0 \leq \varsigma_\ell \leq 1$ for all ℓ.

Using the hazard function recursion [3.1], the sequence $\{\varsigma_\ell\}_{\ell=1}^{\infty}$ satisfies

$$\varsigma_\ell = \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i\right) - \sum_{i=1}^{\min\{\ell-1,n\}} \frac{\varphi_i}{\prod_{j=\ell-i}^{\ell-1} \varsigma_j}, \qquad \text{[A.6.1]}$$

for all ℓ.

Consider the claim

$$\sum_{j=i}^{n} \varphi_j \leq \alpha. \qquad \text{[A.6.2]}$$

Since [A.6.1] implies $\varsigma_1 = 1 - \alpha + \sum_{i=1}^{n} \varphi_i$, the requirement $\varsigma_1 \leq 1$ implies that [A.6.2] is true for $i = 1$.

Now suppose that the claim [A.6.2] has been proved for all $i = 1, \ldots, k$ for some k. If $\varphi_k \geq 0$ then the result $\sum_{j=k+1}^{n} \varphi_j \leq \alpha$ follows automatically from $\sum_{j=k}^{n} \varphi_j \leq \alpha$, proving the statement [A.6.2] for the case $i = k+1$ as well.

Consider the case $\varphi_k < 0$. Using [A.6.1], the requirement $\varsigma_{k+1} \leq 1$ is equivalent to

$$-\sum_{i=1}^{k-1} \frac{\varphi_i}{\prod_{j=k+1-i}^{k} \varsigma_j} - \frac{\varphi_k}{\prod_{j=1}^{k} \varsigma_j} \leq \alpha - \sum_{i=1}^{n} \varphi_i. \quad \text{[A.6.3]}$$

Since $0 \leq \varsigma_1 \leq 1$ and $\varphi_k < 0$ in the case under consideration, it follows from [A.6.3] that

$$-\sum_{i=1}^{k-2} \frac{\varphi_i}{\prod_{j=k+1-i}^{k} \varsigma_j} - \frac{(\varphi_{k-1} + \varphi_k)}{\prod_{j=2}^{k} \varsigma_j} \leq \alpha - \sum_{i=1}^{n} \varphi_i. \quad \text{[A.6.4]}$$

Now if $\varphi_{k-1} + \varphi_k \geq 0$ then $\sum_{j=k+1}^{n} \varphi_j \leq \alpha$ would follow from $\sum_{j=k-1}^{n} \varphi_j \leq \alpha$, proving the statement [A.6.2] for $i = k+1$. If not, then since $0 \leq \varsigma_2 \leq 1$, inequality [A.6.4] together with $\varphi_{k-1} + \varphi_k < 0$ implies that

$$-\sum_{i=1}^{k-3} \frac{\varphi_i}{\prod_{j=k+1-i}^{k} \varsigma_j} - \frac{(\varphi_{k-2} + \varphi_{k-1} + \varphi_k)}{\prod_{j=3}^{k} \varsigma_j} \leq \alpha - \sum_{i=1}^{n} \varphi_i. \quad \text{[A.6.5]}$$

By again considering the two cases for the sign of $\varphi_{k-2} + \varphi_{k-1} + \varphi_k$ the claim [A.6.2] for $i = k+1$ either follows, or a new inequality is deduced alone the pattern of [A.6.3]–[A.6.5] above. This process terminates either with [A.6.2] proved for $i = k+1$ or the inequality

$$-\frac{\sum_{i=1}^{k} \varphi_i}{\varsigma_k} \leq \alpha - \sum_{i=1}^{n} \varphi_i.$$

Since $0 \leq \varsigma_k \leq 1$ and the claim [A.6.2] is known to be true for $i = 1$, it follows that

$$-\sum_{i=1}^{k} \varphi_i \leq \alpha - \sum_{i=1}^{n} \varphi_i,$$

which proves that [A.6.2] holds for $i = k+1$. Thus, [A.6.2] is true for $i = k+1$ in all cases, so it follows for all $i = 1, \ldots, n$ by induction.

Next, consider the claim

$$-(1-\alpha) \leq \sum_{j=i}^{n} \varphi_i. \quad \text{[A.6.6]}$$

Noting that [A.6.1] implies $\varsigma_1 = 1 - \alpha + \sum_{i=1}^{n} \varphi_i$, the requirement $\varsigma_1 \geq 0$ means that [A.6.6] must hold for $i = 1$.

Now suppose that the statement [A.6.6] has been proved for $i = 1, \ldots, k$ for some k. Given equation [A.6.1], the inequality $\varsigma_{k+1} \geq 0$ holds if and only if

$$\sum_{i=1}^{k} \frac{\varphi_i}{\prod_{j=k+1-i}^{k} \varsigma_j} \leq \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i\right). \quad \text{[A.6.7]}$$

Multiplying both sides by the non-negative term $\prod_{j=1}^{k} \varsigma_j$ leads to an equivalent inequality:

$$\sum_{i=1}^{k-1} \left(\prod_{j=1}^{k-i} \varsigma_j\right) \varphi_i + \varphi_k \leq \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i\right) \prod_{j=1}^{k} \varsigma_j. \quad \text{[A.6.8]}$$

If $\varphi_k < 0$ then the inequality $-(1-\alpha) \leq \sum_{j=k+1}^{n} \varphi_j$ follows automatically from $-(1-\alpha) \leq \sum_{j=k}^{n} \varphi_j$, proving the statement [A.6.6] for $i = k+1$. On the other hand, if $\varphi_k \leq 0$ then inequality [A.6.8] together with the requirement $0 \leq \varsigma_1 \leq 1$ implies

$$\sum_{i=1}^{k-2} \left(\prod_{j=1}^{k-i} \varsigma_j \right) \varphi_i + \varsigma_1 (\varphi_{k-1} + \varphi_k) \leq \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i \right) \prod_{j=1}^{k} \varsigma_j. \qquad \text{[A.6.9]}$$

If $\varphi_{k-1} + \varphi_k < 0$ then $-(1-\alpha) \leq \sum_{j=k+1}^{n} \varphi_j$ follows from knowing $-(1-\alpha) \leq \sum_{j=k-1}^{n} \varphi_j$, proving [A.6.6] for $i = k+1$. But if $\varphi_{k-1} + \varphi_k \geq 0$ then [A.6.9] and $0 \leq \varsigma_2 \leq 1$ imply:

$$\sum_{i=1}^{k-3} \left(\prod_{j=1}^{k-i} \varsigma_j \right) \varphi_i + \varsigma_1 \varsigma_2 (\varphi_{k-2} + \varphi_{k-1} + \varphi_k) \leq \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i \right) \prod_{j=1}^{k} \varsigma_j. \qquad \text{[A.6.10]}$$

Proceeding this way, the claim [A.6.6] either follows, or the following inequality is eventually deduced:

$$\left(\prod_{j=1}^{k-1} \varsigma_j \right) \left(\sum_{i=1}^{k} \varphi_i \right) \leq \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i \right) \prod_{j=1}^{k} \varsigma_j.$$

Since [A.6.6] is known to be true for $i = 1$ and as $0 \leq \varsigma_k \leq 1$, it follows that:

$$\left(\prod_{j=1}^{k-1} \varsigma_j \right) \left(\sum_{i=1}^{k} \varphi_i \right) \leq \left(1 - \alpha + \sum_{i=1}^{n} \varphi_i \right) \left(\prod_{j=1}^{k-1} \varsigma_j \right),$$

from which the statement [A.6.6] is proved for $i = k+1$. Thus [A.6.6] is demonstrated for all $i = 1, \ldots, n$ by induction. Therefore $-(1-\alpha) \leq \sum_{j=i}^{n} \varphi_j \leq \alpha$ for all $i = 1, \ldots, n$.

(ii) Suppose $n = 1$ and $\varphi \equiv \varphi_1$. In the case $\varphi = 0$, the restriction $0 \leq \alpha \leq 1$ is clearly all that is required for the hazard function to be well defined. Thus assume $\varphi \neq 0$ in what follows.

The hazard function recursion [3.1] in the case $n = 1$ reduces to

$$\alpha_\ell = (\alpha - \varphi) + \frac{\varphi}{1 - \alpha_{\ell-1}}, \qquad \text{[A.6.11]}$$

and the linear recursion for the survival probabilities [3.2] becomes:

$$\psi_\ell = (1 - \alpha + \varphi)\psi_{\ell-1} - \varphi \psi_{\ell-2}. \qquad \text{[A.6.12]}$$

Define the quadratic equation $\phi(z) = 1 - (1 - \alpha + \varphi)z - \varphi z^2$. Note that $\phi(z^{-1}) = 0$ is the characteristic equation for the sequence of survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$. Let ζ_1 and ζ_2 denote the reciprocals of the two roots of $\phi(z) = 0$. The quadratic can thus be written as $\phi(z) = (1 - \zeta_1 z)(1 - \zeta_2 z)$. By equating coefficients of powers of z, it follows that $1 - \alpha + \varphi = \zeta_1 + \zeta_2$ and $\varphi = \zeta_1 \zeta_2$. Note that $\alpha_1 = \alpha - \varphi$, which must be a well-defined probability, so $\varphi \leq \alpha$ is always required.

The roots ζ_1 and ζ_2 are real numbers when the following condition is satisfied:

$$(1 - \alpha + \varphi)^2 - 4\varphi = \varphi^2 - 2(1 + \alpha)\varphi + (1 - \alpha)^2 \geq 0. \qquad \text{[A.6.13]}$$

Interpreted as a quadratic in φ, it is straightforward to see that it has two positive real roots. The condition above is satisfied when φ is below the smaller of the two roots:

$$\varphi \leq (1 + \alpha) - \sqrt{(1 + \alpha)^2 - (1 - \alpha)^2} = \left(1 - \sqrt{\alpha} \right)^2. \qquad \text{[A.6.14]}$$

The sum of the roots of the quadratic in [A.6.13] is $2(1 + \alpha)$, so the larger root is greater than α, which is in the range where $\varphi \leq \alpha$ is violated.

Consider first the case where $\varphi > 0$. Suppose it is claimed that there is an upper bound $\bar{\alpha}$ for the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$. If $\alpha_{\ell-1} \leq \bar{\alpha}$ then [A.6.11] implies

$$\alpha_\ell \leq (\alpha - \varphi) + \frac{\varphi}{1-\bar{\alpha}}.$$

Hence $\bar{\alpha}$ is valid upper bound for $\{\alpha_\ell\}_{\ell=1}^{\infty}$ (satisfying $0 < \bar{\alpha} < 1$) if the following inequality holds:

$$(\alpha - \varphi) + \frac{\varphi}{1-\bar{\alpha}} \leq \bar{\alpha},$$

which is equivalent to:

$$1 - (1 - \alpha + \varphi)(1-\bar{\alpha})^{-1} + \varphi(1-\bar{\alpha})^{-2} \leq 0.$$

Since in the case $\varphi > 0$, [A.6.11] implies the hazard function is strictly increasing as long as it remains well defined. Thus the hazard function is well defined if and only if $\varphi \leq \alpha$ and there is some bound $\bar{\alpha}$ satisfying $0 < \bar{\alpha} < 1$ such that $\phi((1-\bar{\alpha})^{-1}) \leq 0$. This requires $\phi(z) = 0$ to have real roots, which in turn requires the inequality in [A.6.14] to be satisfied. Furthermore, one of the real roots must be strictly greater than one to ensure $0 < \bar{\alpha} < 1$. Note that $\phi(0) = 1$ and $\phi(1) = \alpha > 0$, and that the product of the roots of $\phi(z) = 0$ is φ^{-1}. Under the condition [A.6.14], $\varphi < 1$, so the product of the roots is greater than one. The sum of the roots is positive, so both must be positive. Thus, $\varphi \leq \alpha$ and [A.6.14] are necessary and sufficient for the hazard function to be well defined in the case $\varphi > 0$.

Now consider the case where $\varphi < 0$. Since $\alpha_1 = \alpha - \varphi$, it is necessary to assume $\varphi \geq -(1-\alpha)$ to ensure α_1 is a well-defined probability. Note that any negative value of φ satisfies [A.6.13], so both ζ_1 and ζ_2 are real numbers. As $\zeta_1\zeta_2 = \varphi$, one of these numbers must be positive and the other negative. Without loss of generality, assume $\zeta_1 > 0$ and $\zeta_2 < 0$. Since $\zeta_1 + \zeta_2 = 1 - \alpha + \varphi$ and as $\alpha_1 = \alpha - \varphi$ is well defined, it follows that $\zeta_1 > -\zeta_2$. Noting that $\phi(0) = 1$ and $\phi(1) = \alpha$, so as $\phi(\zeta_1^{-1}) = 0$ and $\phi(\zeta_2^{-1}) = 0$ it must be the case that $\zeta_1 < 1$ (otherwise $\phi(z)$ would have to change sign twice between 0 and 1, implying that both ζ_1 and ζ_2 would be positive).

Since ζ_1 and ζ_2 are distinct numbers in the case $\varphi < 0$, the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ can be expressed as $\psi_\ell = \varkappa_1\zeta_1^\ell + \varkappa_2\zeta_2^\ell$, where $\varkappa_1$ and $\varkappa_2$ are real numbers. Consequently:

$$\psi_\ell = \varkappa_1\zeta_1^\ell\left\{1 + \frac{\varkappa_2}{\varkappa_1}\left(\frac{\zeta_2}{\zeta_1}\right)^\ell\right\}, \quad \text{and} \quad \psi_\ell - \psi_{\ell+1} = \varkappa_1(1-\zeta_1)\zeta_1^\ell\left\{1 + \frac{\varkappa_2(1-\zeta_2)}{\varkappa_1(1-\zeta_1)}\left(\frac{\zeta_2}{\zeta_1}\right)^\ell\right\}. \qquad \text{[A.6.15]}$$

The hazard function recursion [A.6.11] implies $\alpha_1 = \alpha - \varphi$ and $\alpha_2 = (\alpha - \varphi) + \varphi/(1 - \alpha + \varphi)$. Given the restriction $\varphi \geq -(1-\alpha)$ that ensures α_1 is well defined in the case $\varphi < 0$, the probability α_2 is well defined if and only if $\varphi \geq -(\alpha - \varphi)(1 - \alpha + \varphi)$. Rearranging this inequality shows that it is equivalent to

$$\varphi^2 - 2\alpha\varphi - \alpha(1-\alpha) \leq 0.$$

Interpreted as a quadratic in φ, the above inequality has one positive and one negative root. Given that $\varphi < 0$ in the case under consideration, the relevant restriction is that

$$\varphi \geq \alpha - \sqrt{\alpha^2 + \alpha(1-\alpha)} = -\sqrt{\alpha}(1-\sqrt{\alpha}). \qquad \text{[A.6.16]}$$

Notice that $\sqrt{\alpha}(1-\sqrt{\alpha}) \leq 1 - \alpha$, so the requirement $\varphi \geq -(1-\alpha)$ is automatically satisfied when [A.6.16] holds.

The condition [A.6.16] is thus seen to be equivalent to α_1 and α_2 being well defined in the case $\varphi < 0$. This is itself equivalent to $0 \leq \psi_2 \leq \psi_1 \leq \psi_0 = 1$ because $\psi_\ell = (1-\alpha_\ell)\psi_{\ell-1}$. By using [A.6.15], $\psi_0 - \psi_1 = \varkappa_1(1-\zeta_1) + \varkappa_2(1-\zeta_2) \geq 0$ and $\psi_1 - \psi_2 = \varkappa_1(1-\zeta_1)\zeta_1 + \varkappa_2(1-\zeta_2)\zeta_2 \geq 0$. Since $0 < \zeta_1 < 1$ and $\zeta_2 < 0$, it follows from the first inequality that at least one of $\varkappa_1$ and $\varkappa_2$ must be non-negative, and thus from the second inequality that $\varkappa_1 \geq 0$.

Since $\zeta_1 > -\zeta_2$, the terms $(\varkappa_2/\varkappa_1)(\zeta_2/\zeta_1)^i$ and $(\varkappa_2(1-\zeta_2)/\varkappa_1(1-\zeta_1))(\zeta_2/\zeta_1)^i$ in [A.6.15] must alternate in sign and decline in absolute value as ℓ increases. Because $\varkappa_1$, ζ_1 and $(1-\zeta_1)$ are non-negative, the inequalities $0 \leq \psi_2 \leq \psi_1 \leq \psi_0$ imply $0 \leq \psi_\ell \leq \psi_{\ell-1}$ for all ℓ, which ensure the hazard function is well defined everywhere. Since this condition is equivalent to [A.6.16], the proof is complete.

A.7 Proof of Proposition 6

(i) Let $\psi(z) \equiv \sum_{\ell=0}^{\infty} \psi_\ell z^\ell$ and $\omega(z) \equiv \sum_{\ell=0}^{\infty} \omega_\ell z^\ell$ denote the z-transforms of the survival probabilities $\{\psi_\ell\}_{\ell=0}^{\infty}$ and the stationary age distribution $\{\omega_\ell\}_{\ell=0}^{\infty}$. Equations [2.4] and [2.5] for the reset price r_t and price level p_t can be written in terms of the lag and forward operators $\mathbb{L}$ and $\mathbb{F}$ and the power series $\psi(z)$ and $\omega(z)$:

$$\mathsf{r}_t = \psi(\beta)^{-1}\mathbb{E}_t\left[\psi(\beta\mathbb{F})\mathsf{p}_t^*\right], \quad \text{and} \quad \mathsf{p}_t = \omega(\mathbb{L})\mathsf{r}_t. \qquad \text{[A.7.1]}$$

Suppose that the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ is generated by the recursion [3.1] using parameters α and $\{\varphi_i\}_{i=1}^{n}$. Define the polynomial $\phi(z) \equiv 1 - (1 - \alpha + \sum_{i=1}^{n} \varphi_i)\, z + \sum_{j=1}^{n} \varphi_j z^{j+1}$. Lemma 1 shows that $\psi(z)$ is analytic on the set $\mathscr{D}_\rho \equiv \left\{z \in \mathbb{C} \;\middle|\; |z| \leq \rho\right\}$ for some $\rho > 1$. Note that the equivalent recursion [3.2] for the survival probabilities and $\psi_0 = 1$ imply $\phi(z)\psi(z) = 1$ for all $z \in \mathscr{D}_\rho$.

Now multiply both sides of the equation in [A.7.1] for the reset price r_t by $\phi(\beta\mathbb{F})$ and take conditional expectations at time t:

$$\mathbb{E}_t\left[\phi(\beta\mathbb{F})\mathsf{r}_t\right] = \mathbb{E}_t\left[\phi(\beta\mathbb{F})\left\{\psi(\beta)^{-1}\mathbb{E}_t\left[\psi(\beta\mathbb{F})\mathsf{p}_t^*\right]\right\}\right] = \mathbb{E}_t\left[\psi(\beta)^{-1}\mathbb{E}_t\left[\phi(\beta\mathbb{F})\psi(\beta\mathbb{L})\mathsf{p}_t^*\right]\right] = \psi(\beta)^{-1}\mathsf{p}_t^*. \qquad \text{[A.7.2]}$$

This result follows first because $\phi(\beta\mathbb{F})$ contains only non-negative powers of $\mathbb{F}$, so it commutes with the conditional expectation $\mathbb{E}_t[\cdot]$ operator inside another conditional expectation. Second, $\phi(z)\psi(z) = 1$, hence $\phi(\beta\mathbb{F})\psi(\beta\mathbb{F}) = \mathbb{I}$, where $\mathbb{I}$ is the identity operator. Next, note that because $\omega_\ell = (1 - \alpha_\ell)\omega_{\ell-1}$ and $\psi_\ell = (1 - \alpha_\ell)\psi_{\ell-1}$, the functions $\omega(z)$ and $\psi(z)$ are proportional. Thus $\omega(z) = (\omega(1)/\psi(1))\psi(z)$, and $\omega(z) = \phi(1)\psi(z)$, since $\psi(1)^{-1} = \phi(1)$, and $\omega(1) = 1$ because $\{\omega_\ell\}_{\ell=0}^{\infty}$ is a probability distribution. It follows that $\phi(z)\omega(z) = \phi(1)$ for all $z \in \mathscr{D}_\rho$. Multiplying both sides of the equation for p_t in [A.7.1] by $\phi(\mathbb{L})$ yields

$$\phi(\mathbb{L})\mathsf{p}_t = \phi(\mathbb{L})\omega(\mathbb{L})\mathsf{r}_t = \phi(1)\mathbb{I}\mathsf{r}_t = \phi(1)\mathsf{r}_t. \qquad \text{[A.7.3]}$$

Now multiply both sides of equation [A.7.2] by $\phi(1)$ and note that $\psi(\beta)^{-1} = \phi(\beta)$, and then substitute the expression for p_t^* from [2.3]:

$$\mathbb{E}_t\left[\phi(\beta\mathbb{F})\phi(1)\mathsf{r}_t\right] = \phi(1)\phi(\beta)(\mathsf{p}_t + \nu\mathsf{x}_t).$$

Substitute the formula for $\phi(1)\mathsf{r}_t$ from [A.7.3] into the above and divide both sides by $\phi(1)\phi(\beta)$:

$$\mathbb{E}_t\left[\left\{\frac{\phi(\mathbb{L})}{\phi(1)}\frac{\phi(\beta\mathbb{F})}{\phi(\beta)} - 1\right\}\mathsf{p}_t\right] = \nu\mathsf{x}_t. \qquad \text{[A.7.4]}$$

Define the Laurent polynomial $\chi(z)$ as follows:

$$\chi(z) \equiv \frac{\phi(z)}{\phi(1)}\frac{\phi(\beta z^{-1})}{\phi(\beta)} - 1, \qquad \text{[A.7.5]}$$

so that equation [A.7.4] is equivalent to $\mathbb{E}_t\left[\chi(\mathbb{L})\mathsf{p}_t\right] = \nu\mathsf{x}_t$, noting that $\mathbb{F} \equiv \mathbb{L}^{-1}$. For algebraic convenience, define the sequence of coefficients $\{\phi_j\}_{j=1}^{n+1}$ by $\phi_1 \equiv (1 - \alpha + \sum_{i=1}^{n} \varphi_i)$ and $\phi_j \equiv -\varphi_{j-1}$ for $j = 2, \ldots, n+1$ in terms of the parameters of the recursion [3.1]. With these definitions the polynomial $\phi(z)$ can be written as $\phi(z) \equiv 1 - \sum_{j=1}^{n+1} \phi_j z^j$. The Laurent polynomial $\chi(z)$ can be written explicitly using this expression:

$$\chi(z) = \vartheta\left\{\left(1 - \sum_{j=1}^{n+1} \phi_j z^j\right)\left(1 - \sum_{j=1}^{n+1} \beta^j \phi_j z^{-j}\right) - \left(1 - \sum_{j=1}^{n+1} \phi_j\right)\left(1 - \sum_{j=1}^{n+1} \beta^j \phi_j\right)\right\},$$

where $\vartheta \equiv \phi(1)^{-1}\phi(\beta)^{-1}$ is defined. Expanding the brackets to obtain an expression of the form $\chi(z) =$

$\sum_{\ell=-(n+1)}^{n+1} \chi_\ell z^\ell$ and equating powers of z implies that $\chi(z)$ can be written as

$$\chi(z) = \chi_0 + \sum_{\ell=1}^{n+1} \chi_\ell \left\{ z^\ell + \beta^\ell z^{-\ell} \right\}, \quad \text{where } \chi_\ell = -\vartheta \left\{ \phi_\ell - \sum_{\jmath=1}^{n+1-\ell} \beta^\jmath \phi_\jmath \phi_{\jmath+\ell} \right\} \text{ for } \ell \geq 1, \tag{A.7.6}$$

since $\chi_{-\ell} = \beta^\ell \chi_\ell$ for all ℓ. As the definition in [A.7.5] implies $\chi(1) = 0$, it follows that $\chi_0 = -\sum_{\ell=1}^{n+1}(1 + \beta^\ell)\chi_\ell$. Furthermore, $\chi(1) = 0$ implies that there exists a Laurent polynomial $\theta(z)$ such that $\chi(z) = (1 - z)\theta(z)$. Given the degree of $\chi(z)$, this Laurent polynomial must have the form $\theta(z) = \sum_{\ell=-(n+1)}^{n} \theta_\ell z^\ell$. Multiplying $\theta(z)$ by $1 - z$ and equating powers of z yields an expression for $\chi(z)$:

$$\chi(z) = \theta_{-(n+1)} z^{-(n+1)} + \sum_{\ell=-n}^{n} (\theta_\ell - \theta_{\ell-1}) z^\ell - \theta_n z^{n+1}.$$

Equating coefficients of powers of z with those in [A.7.6] implies $\chi_{n+1} = -\theta_n$, $\beta^{n+1}\chi_{n+1} = \theta_{-(n+1)}$, and $\chi_\ell = \theta_\ell - \theta_{\ell-1}$ for all $\ell = -n, \ldots, n$. Iterating these relationships then implies

$$\theta_\ell = -\sum_{\jmath=\ell+1}^{n+1} \chi_\jmath, \quad \text{and } \theta_{-\ell} = \sum_{\jmath=\ell}^{n+1} \beta^\jmath \chi_\jmath, \tag{A.7.7}$$

for all $\ell = 1, \ldots, n+1$. Combining these expressions with those for χ_i in [A.7.6] yields

$$\theta_\ell = \vartheta \sum_{i=\ell+1}^{n+1} \left\{ \phi_i - \sum_{j=1}^{n+1-i} \beta^j \phi_j \phi_{i+j} \right\} = \vartheta \sum_{i=\ell+1}^{n+1} \phi_i \left\{ 1 - \sum_{j=1}^{i-\ell-1} \beta^j \phi_j \right\}, \tag{A.7.8a}$$

where a change in the order of summation has been made in the final term. Similarly,

$$\theta_{-\ell} = -\vartheta \sum_{i=\ell}^{n+1} \beta^i \left\{ \phi_i - \sum_{j=1}^{n+1-i} \beta^j \phi_j \phi_{i+j} \right\} = -\vartheta \sum_{i=\ell}^{n+1} \beta^i \phi_i \left\{ 1 - \sum_{j=1}^{i-\ell} \phi_j \right\}. \tag{A.7.8b}$$

The original definitions of the terms of the sequence $\{\phi_\jmath\}_{\jmath=1}^{n+1}$ are $\phi_1 = 1 - \alpha + \sum_{i=1}^{n} \varphi_i$ and $\phi_\jmath = -\varphi_{\jmath-1}$ for $\jmath = 2, \ldots, n+1$. Substituting the original parameters α and $\{\varphi_\jmath\}_{\jmath=1}^{n}$ back into [A.7.8a] and [A.7.8b] yields

$$\theta_\ell = -\vartheta \left\{ \varphi_\ell + \sum_{i=\ell+1}^{n} \varphi_i \left(1 - \beta(1 - \alpha_1) + \sum_{j=1}^{i-\ell-1} \beta^{j+1} \varphi_j \right) \right\}, \quad \text{for } \ell = 1, \ldots, n; \tag{A.7.9a}$$

$$\theta_{-(\ell+1)} = \vartheta \beta^{\ell+1} \left\{ \varphi_\ell + \sum_{i=\ell+1}^{n} \beta^{i-\ell} \varphi_i \left(\alpha_1 + \sum_{j=1}^{i-\ell-1} \varphi_j \right) \right\}, \quad \text{for } \ell = 1, \ldots, n; \tag{A.7.9b}$$

$$\theta_0 = \vartheta \left\{ (1 - \alpha_1) - \sum_{i=1}^{n} \varphi_i \left(1 - \beta(1 - \alpha_1) + \sum_{j=1}^{i-1} \beta^{j+1} \varphi_j \right) \right\}; \text{ and} \tag{A.7.9c}$$

$$\theta_{-1} = -\vartheta \beta \left\{ (1 - \alpha_1) - \sum_{i=1}^{n} \beta^i \varphi_i \left(\alpha_1 + \sum_{j=1}^{i-1} \varphi_j \right) \right\}. \tag{A.7.9d}$$

Since the definition of $\theta(z)$ requires $\chi(z) = (1 - z)\theta(z)$, and as inflation is defined by $\pi_t = (\mathbb{I} - \mathbb{L})\mathsf{p}_t$, it follows from equations [A.7.4] and [A.7.5] that $\mathbb{E}_t[\theta(\mathbb{L})\pi_t] = \nu x_t$. Make the following definitions of the coefficient κ and the sequences $\{\lambda_\ell\}_{\ell=1}^{n}$ and $\{\xi_\ell\}_{\ell=1}^{n+1}$ in terms of the elements of the sequence $\{\theta_\ell\}_{\ell=-(n+1)}^{n}$

from [A.7.9]:

$$\lambda_\ell \equiv -\frac{\theta_\ell}{\theta_0}, \quad \xi_\ell \equiv -\frac{\theta_{-\ell}}{\theta_0}, \quad \text{and} \quad \kappa \equiv \frac{1}{\theta_0}, \tag{A.7.10}$$

noting that $\theta_0 > 0$ ensures these definitions are valid. With these definition, the Laurent polynomial $\theta(z)$ is given by $\theta(z) = \kappa^{-1}\left\{1 - \sum_{\ell=1}^{n}\lambda_\ell z^\ell - \sum_{\ell=1}^{n+1}\xi_\ell z^{-\ell}\right\}$, and so $\mathbb{E}_t\left[\theta(\mathbb{L})\pi_t\right] = \nu x_t$ is equivalent to the Phillips curve in [4.3].

(ii) First consider the expression for θ_0 in [A.7.9c]. By expanding the bracket and changing the order of summation:

$$\theta_0 = \vartheta\left\{(1-\alpha_1) - (1-\beta(1-\alpha_1))\left(\sum_{i=1}^{n}\varphi_i\right) - \sum_{i=1}^{n-1}\beta^{i+1}\varphi_i\left(\sum_{j=i+1}^{n}\varphi_j\right)\right\}.$$

Adding and subtracting terms in the final summation to obtain an equivalent expression:

$$\theta_0 = \vartheta\left\{(1-\alpha_1) - (1-\beta(1-\alpha_1))\left(\sum_{i=1}^{n}\varphi_i\right) - \left(\sum_{i=1}^{n}\beta^{i+1}\varphi_i\right)\left(\sum_{j=1}^{n}\varphi_j\right) + \sum_{i=1}^{n}\beta^{i+1}\varphi_i\left(\sum_{j=1}^{i}\varphi_j\right)\right\}.$$

The definition of the polynomial $\phi(z)$ implies $\phi(1) = \alpha_1 + \sum_{i=1}^{n}\varphi_i$ and $\phi(\beta) = 1 - \beta(1-\alpha_1) + \sum_{i=1}^{n}\beta^{i+1}\varphi_i$. By defining the sums $s_i \equiv \sum_{j=1}^{i}\varphi_j$ for $i = 0, \ldots, n$ (with $s_0 = 0$) and noting that $\phi(1) - \alpha_1 = \sum_{i=1}^{n}\varphi_i$:

$$\theta_0 = \vartheta\left\{(1-\alpha_1) - (\phi(1)-\alpha_1)\,\phi(\beta) + \sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1})s_i\right\}.$$

Rearranging the first two terms leads to

$$\theta_0 = \vartheta\left\{(1-\alpha_1)(1-\phi(\beta)) + (1-\phi(1))\phi(\beta) + \sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1})s_i\right\}. \tag{A.7.11}$$

Note that $(s_i - s_{i-1})s_i = (1/2)\left\{(s_i^2 - s_{i-1}^2) + (s_i - s_{i-1})^2\right\}$, and thus

$$\sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1})s_i = \frac{1}{2}\left\{\sum_{i=1}^{n}\beta^{i+1}\varphi_i^2 + (1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i^2 + \beta^{n+1}s_n^2 - \beta^2 s_0^2\right\},$$

since $s_i - s_{i-1} = \varphi_i$. Using $s_0 = 0$ and substituting this result into [A.7.11]:

$$\theta_0 = \vartheta\left\{(1-\alpha_1)(1-\phi(\beta)) + (1-\phi(1))\phi(\beta) + \frac{1}{2}\left\{\sum_{i=1}^{n}\beta^{i+1}\varphi_i^2 + (1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i^2 + \beta^{n+1}s_n^2\right\}\right\}. \tag{A.7.12}$$

Now observe that $\phi(1) = \psi(1)^{-1}$, and $\psi(1) = \sum_{\ell=0}^{\infty}\psi_\ell$, so $0 < \phi(1) < 1$ because $\psi_0 \equiv 1$, $\psi_\ell \geq 0$, $\sum_{\ell=0}^{\infty}\psi_\ell < \infty$, and $\psi_1 > 0$ under the assumption $\alpha_1 < 1$. Similarly, $\phi(\beta) = \psi(\beta)^{-1}$ and $\psi(\beta) = \sum_{\ell=0}^{\infty}\beta^\ell\psi_\ell$. Since $0 < \beta < 1$, it follows that $0 < \phi(\beta) < 1$. Together these results establish that $\vartheta > 0$ since $\vartheta \equiv \phi(1)^{-1}\phi(\beta)^{-1}$. Given $\alpha_1 < 1$, the parameter α_1 must satisfy $0 \leq \alpha_1 < 1$. Consequently, the first two terms in the brackets in [A.7.12] are strictly positive and all other terms are non-negative. Thus, it is shown that $\theta_0 > 0$. The proof of $\theta_0 > 0$ then automatically shows $\kappa > 0$.

(iii) Now consider the value of ξ_1, which requires examining θ_{-1}. Let $s_i \equiv \alpha_1 + \sum_{j=1}^{i}\varphi_j$, and so $s_i - s_{i-1} = \varphi_i$ for all $i = 1, \ldots, n$, and $s_0 = \alpha_1$. The expression for θ_{-1} in [A.7.9d] can be written as

$$-\theta_{-1} = \vartheta\left\{\beta(1-\alpha_1) - \sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1})s_{i-1}\right\}. \tag{A.7.13}$$

Note that $(s_i - s_{i-1})s_{i-1} = (1/2)\left\{(s_i^2 - s_{i-1}^2) - (s_i - s_{i-1})^2\right\}$, and hence

$$\sum_{i=1}^{n} \beta^{i+1}(s_i - s_{i-1})s_{i-1} = \frac{1}{2}\left\{(1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i^2 + \beta^{n+1}s_n^2 - \beta^2 s_0^2 - \sum_{i=1}^{n}\beta^{i+1}\varphi_i^2\right\}.$$

Also note that $\sum_{i=1}^{n}\beta^{i+1}\varphi_i = \sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1}) = (1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i + \beta^{n+1}s_n - \beta^2 s_0$. By adding and subtracting a multiple of these equal terms to the equation above:

$$\sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1})s_{i-1} = \frac{1}{2}\sum_{i=1}^{n}\beta^{i+1}\varphi_i - \frac{1}{2}\left\{(1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i + \beta^{n+1}s_n - \frac{1}{2}\beta^2\alpha_1\right\}$$
$$+\frac{1}{2}\left\{(1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i^2 + \beta^{n+1}s_n^2 - \beta^2\alpha_1^2 - \sum_{i=1}^{n}\beta^{i+1}\varphi_i^2\right\},$$

recalling that $s_0 = \alpha_1$. Since $\sum_{i=1}^{\infty}\beta^{i+1}\varphi_i = \phi(\beta) - 1 + \beta(1-\alpha_1)$, the above equation can be rearranged as follows:

$$\sum_{i=1}^{n}\beta^{i+1}(s_i - s_{i-1})s_{i-1} = -\frac{\beta}{2}\left\{\alpha_1^2\beta - \alpha_1\beta + \alpha_1 - 1\right\} + \frac{1}{2}(1-\phi(\beta))$$
$$+\frac{1}{2}\left\{(1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i(1-s_i) + \beta^{n+1}s_n(1-s_n) + \sum_{i=1}^{n}\beta^{i+1}\varphi_i^2\right\}.$$

Substituting this result into equation [A.7.13] yields:

$$-\theta_{-1} = \frac{\vartheta}{2}\left\{\beta(1-\alpha_1)(1-\alpha_1\beta) + (1-\phi(\beta))\right\}$$
$$+\frac{\vartheta}{2}\left\{(1-\beta)\sum_{i=1}^{n-1}\beta^{i+1}s_i(1-s_i) + \beta^{n+1}s_n(1-s_n) + \sum_{i=1}^{n}\beta^{i+1}\varphi_i^2\right\}.$$

Proposition 5 demonstrates that $0 \leq s_i \leq 1$ for all $i = 1, \ldots, n$ is necessary for the hazard function $\{\alpha_\ell\}_{\ell=1}^{\infty}$ to be well defined. Since $0 < \alpha_1 < 1$, $0 < \beta < 1$, and $0 < \phi(\beta) < 1$, $\vartheta > 0$ and $\theta_0 > 0$ as shown earlier, it follows that $\xi_1 = -\theta_{-1}/\theta_0$ is strictly positive.

(iv) Next, note that

$$1 - \beta(1-\alpha_1) + \sum_{j=1}^{i}\beta^{j+1}\varphi_j = (1-\beta)\left\{1 + \sum_{h=0}^{i-1}\beta^{h+1}s_h\right\} + \beta^{i+1}s_i, \quad \text{where } s_j \equiv \alpha_1 + \sum_{h=1}^{j}\varphi_j.$$

Using equations [A.7.9a] and [A.7.10], the coefficients of lagged inflation $\{\lambda_\ell\}_{\ell=1}^{n}$ can be expressed as

$$\lambda_\ell = \left(\frac{\vartheta}{\theta_0}\right)\varphi_\ell + \sum_{i=\ell+1}^{n}\left\{\left(\frac{\vartheta}{\theta_0}\right)\left((1-\beta)\left(1 + \sum_{j=0}^{i-\ell-2}\beta^{j+1}s_j\right) + \beta^{i-\ell}s_{j-i-1}\right)\right\}\varphi_i.$$

Proposition 5 shows that $0 \leq s_i \leq 1$ for all $i = 0, 1, \ldots, n$. Since $\vartheta > 0$ and $\theta_0 > 0$, it follows that λ_ℓ is a weighted sum of $\varphi_\ell, \ldots, \varphi_n$.

(v) Similarly, equations [A.7.9b] and [A.7.10] show that the coefficients on future inflation $\{\xi_\ell\}_{\ell=2}^{n+1}$ are:

$$\xi_\ell = -\left\{\left(\frac{\vartheta\beta^\ell}{\theta_0}\right)\varphi_{\ell-1} + \sum_{i=\ell}^{n}\left(\frac{\vartheta\beta^{i+1}s_{i-\ell}}{\theta_0}\right)\varphi_i\right\}, \quad \text{for } \ell = 2, \ldots, n+1,$$

where s_i is as defined above. Thus ξ_ℓ for $\ell \geq 2$ is the negative of a weighted sum of the parameters

$\varphi_{\ell-1}, \ldots, \varphi_n$.

(vi) Note that [A.7.5] implies $\chi(\beta) = 0$. Since $\chi(z) = (1-z)\theta(z)$, it must be the case that $\theta(\beta) = 0$ also. The definition of $\theta(z)$ then implies that $\sum_{\ell \to -(n+1)}^{n} \beta^\ell \theta_\ell = 0$. The result follows by using [A.7.10].

(vii) Finally, to derive the restrictions across the sequences of coefficients $\{\lambda_\ell\}_{\ell=1}^{n}$ and $\{\xi_\ell\}_{\ell=1}^{n+1}$, use the definition in [A.7.10] and equation [A.7.7] to deduce:

$$(1-\beta)\sum_{i=\ell}^{n} \beta^i \lambda_i = -\frac{(1-\beta)}{\theta_0}\sum_{i=\ell}^{n} \beta^i \theta_i = \frac{(1-\beta)}{\theta_0}\sum_{i=\ell}^{n}\sum_{j=i+1}^{n+1} \beta^i \chi_j = \frac{(1-\beta)}{\theta_0}\sum_{i=\ell+1}^{n+1}\left\{\sum_{j=\ell}^{i-1}\beta^j\right\}\chi_i,$$

using a change in the order of summation to derive the final equality. Using the formula for the geometric sum yields

$$(1-\beta)\sum_{i=\ell}^{n} \beta^i \lambda_i = \frac{1}{\theta_0}\sum_{i=\ell+1}^{n+1}(\beta^\ell - \beta^i)\chi_i. \qquad \text{[A.7.14]}$$

Thus, adding β to the equation above in the case of $\ell = 1$ and substituting for θ_0 using [A.7.7]

$$\beta + (1-\beta)\sum_{i=1}^{n} \beta^i \lambda_i = \frac{1}{\theta_0}\left\{\sum_{i=2}^{n+1}(\beta - \beta^i)\chi_i - \beta\sum_{i=1}^{n+1}\chi_i\right\} = -\frac{1}{\theta_0}\sum_{i=1}^{n+1}\beta^i\chi_i = -\frac{\theta_{-1}}{\theta_0},$$

with the expression for θ_{-1} taken from [A.7.7]. Given the definition in [A.7.10], the equation for ξ_1 is confirmed. Now substract the expression in [A.7.14] for $\ell \geq 2$ from $\beta^\ell \lambda_{\ell-1}$:

$$\beta^\ell \lambda_{\ell-1} - (1-\beta)\sum_{i=\ell}^{n}\beta^i\lambda_i = -\frac{\beta^\ell \theta_{\ell-1}}{\theta_0} - \frac{1}{\theta_0}\sum_{i=\ell+1}^{n+1}(\beta^\ell - \beta^i)\chi_i = \frac{1}{\theta_0}\left\{\beta^\ell\sum_{i=\ell}^{n+1}\chi_i - \sum_{i=\ell+1}^{n+1}(\beta^\ell - \beta^i)\chi_i\right\},$$

making use of equations [A.7.7] and [A.7.10]. It follows that

$$\beta^\ell \lambda_{\ell-1} - (1-\beta)\sum_{i=\ell}^{n}\beta^i\lambda_i = \frac{1}{\theta_0}\sum_{i=\ell}^{n+1}\beta^i\chi_i = \frac{\theta_{-\ell}}{\theta_0},$$

using [A.7.7] again. Therefore the equation for ξ_ℓ is verified for $\ell \geq 2$. This completes the proof.