CENTRE *for* ECONOMIC
PERFORMANCE

CEP Discussion Paper No 837

November 2007

Intrinsic Inflation Persistence

Kevin D. Sheedy

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
It is often argued that the New Keynesian Phillips curve is at odds with the data because it cannot explain inflation persistence — the difficulty of returning inflation immediately to target after a shock without any loss of output. This paper explains how a model where newer prices are stickier than older prices is consistent with this phenomenon, even though it introduces no deviation from optimizing, forwards-looking price setting. The probability of adjusting new and old prices is estimated using a novel method that draws only on macroeconomic data, and the findings strongly support the premise of the model.

JEL Classification: E3
Keywords: inflation persistence, hazard function, time-dependent pricing, New Keynesian Phillips curve

This paper was produced as part of the Centre's Macro Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
I thank Petra Geraats for many helpful discussions. I am grateful to Seppo Honkapohja and Michael Woodford for their comments on an earlier draft of this paper, entitled "Structural Inflation Persistence". The paper has also benefited from the comments of seminar participants at the Bank of England, Birmingham University, Cambridge University, the Federal Reserve Board, the University of Georgia, the London School of Economics, Manchester University, the New York Fed, Universitat Pompeu Fabra, and Warwick University. I thank the Economic and Social Research Council for financial support, and the Journal of Applied Econometrics for an award under its Scholars Programme.

Kevin D. Sheedy is an Associate of the Macro Programme at the Centre for Economic Performance. He is also a lecturer in the Department of Economics, London School of Economics. Email: k.d.sheedy@lse.ac.uk

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 978-0-85328-212-9

1. Introduction

It is often argued that New Keynesian economics cannot explain the persistence of inflation. The New Keynesian Phillips curve (NKPC) predicts that once the factors giving rise to high inflation have passed, inflation can return immediately to target without incurring any loss of output. This surprising and puzzling implication is a consequence of the absence of past inflation rates from the NKPC, making inflation determination a purely forward-looking process.

This paper shows that microfounded models of price stickiness are able to generate intrinsic inflation persistence, defined as inflation inherited from the past that cannot be avoided without suffering a temporary reduction in economic activity. To achieve this, it is necessary to find a theoretical reason why past inflation rates should appear in the Phillips curve with positive coefficients. The key ingredient is that firms are more likely to change older rather than newer prices. This is a plausible pricing strategy when individual prices are costly to adjust and there is base drift in the general price level. Firms are then reluctant to squander resources changing prices that have been posted only recently, when those that have remained fixed for a long time are further from profit-maximizing levels. In contrast to this, the widely used Calvo (1983) price-setting model underlying the NKPC assumes the probability of price adjustment is the same for all prices, irrespective of age.

Intuitively, the existence of intrinsic inflation persistence depends on two opposing forces. To see this, consider the case of a temporary cost-push shock lasting for only one period. When price changes are costly, and with staggering of adjustment times, some firms respond to the shock by raising their prices; others take no immediate action. After the shock has dissipated there are two groups of firms and two countervailing effects on inflation. Those firms that did change price initially now find their relative prices too high and want to reduce their prices in money terms. This is the "roll-back" effect. But since the price level has risen, those firms that did not change price initially now want to raise their money prices to maintain desired relative prices. This is the "catch-up" effect.

When the probability of changing a price is independent of the age of the price, the roll-back effect exactly cancels out the catch-up effect and inflation stabilizes immediately. But when older prices are more likely to be changed than newer prices, catch-up dominates roll-back, resulting in persistent inflation as the price level continues to rise. A policymaker wishing to counteract this persistent inflation must engineer a downturn in economic activity to reduce firms' costs and desired relative prices.

This paper makes three distinct contributions to the literature on inflation and price setting behaviour. First, it demonstrates how the New Keynesian model can be reconciled with intrinsic inflation persistence by relaxing the assumption of a constant probability of price adjustment, an explanation that involves no deviation from the essence of the model since firms remain optimizing and forward looking. Intrinsic persistence is linked to the shape of the hazard function for price changes (the probability of a firm posting a new price as a function of the time since its previous price change). The insight is that upward-sloping hazard functions, where newer prices are stickier than older prices, imply a positive relationship between current and past inflation and hence generate persistent inflation. A downward-sloping hazard function would imply a negative relationship between current and past inflation.

Second, this paper also introduces a methodological innovation in deriving a simple and tractable expression for the Phillips curves implied by wide range of price-setting models with differently shaped hazard functions. These Phillips curves bear a close resemblance to the "hybrid" New Keynesian Phillips curves currently favoured on empirical grounds, but which are widely thought to have weak theoretical foundations. It is then shown that the slope of the hazard function determines whether the coefficients on past inflation are positive or negative. The new expression for the Phillips curve derived here also has applications beyond the scope of this paper. It is straightforward to apply it in situations where the NKPC is currently valued for its ease-of-use, but perhaps not for the realism of its assumptions. Given the ubiquity of the NKPC in modern monetary policy analysis (see for example Woodford (2003)), it is important to be able to work with a more general model of price setting, while retaining much of the NKPC's user-friendliness.[1]

Third, there is an empirical contribution in estimating the hazard function for price changes using only macroeconomic data. It is shown how the hazard function can be identified and estimated without the need to have observations of individual prices. It turns out that this can be achieved with surprisingly simple econometric techniques. Using this approach, it is possible to test whether there exists a hazard function model that is consistent with observed inflation dynamics when firms set prices in a purely forward-looking manner. The estimated hazard function can also be compared to those from the burgeoning microeconometric literature. Estimates of the hazard function are presented using U.S. data, and these provide strong support for a model in which newer prices are stickier than older prices.

The problem of inflation persistence was first brought to the attention of economists by Fuhrer and Moore (1995). There are actually several stylized facts about inflation dynamics documented by Mankiw (2001) that the New Keynesian Phillips curve cannot explain on its own, including the cost of disinflation. The NKPC has also been the subject of direct econometric studies in work such as Galí and Gertler (1999), Sbordone (2002), Rudd and Whelan (2005) and Roberts (2005). The extent to which inflation determination is forward looking as opposed to backward looking is hotly debated, but most studies find that the data support a significant backward-looking component, which is conspicuously absent from the standard New Keynesian model.

Fuhrer and Moore's own solution to this problem is a relative contracting model where nominal wages are set with a concern for achieving a real wage that tracks the real wages of other workers. This creates a role for past inflation, because high inflation is associated with high wage growth that indicates one cohort of workers pulling away from the others. Holding

[1]The methodology is applied by Sheedy (2007) to the question of optimal monetary policy.

other determinants of inflation constant, high inflation in the past makes it more likely that there will be inflationary wage pressure in the future as the other cohorts of workers try to close the gap. A widely used alternative theory states that some fraction of firms relies on a rule of thumb when setting prices (Galí and Gertler, 1999). These firms do not maximize profits when they post a new price, but instead simply take past prices posted by other firms and add on a correction for recent inflation. Another popular explanation is proposed by Christiano, Eichenbaum and Evans (2005), who argue that in between those times when actual pricing decisions are made, firms continually re-index their prices in line with past inflation. A variant of this hypothesis has also been put forward by Smets and Wouters (2003). While it is possible to make a case for each of these ideas, what unites them is an essentially arbitrary role assigned to past inflation in the process of setting prices or wages. Each model resolves the problem of inflation determination being partially backward looking by assuming that at least some agents behave in a backward-looking fashion. This paper takes an alternative approach and argues that inflation can be significantly backward looking even when all agents remain optimizing and forward looking.

Others account for inflation persistence by arguing that it results from inflation expectations not being formed rationally (Paloviita, 2004; Roberts, 1997). In a similar vein, a process of adaptive learning by agents also explains some persistence (Milani, 2005). Furthermore, it might be the case that time-variation in the average rate of inflation generates apparent inflation persistence (Cogley and Sbordone, 2005). The importance of these ideas is discussed further in Woodford (2007).

Pricing models with non-constant hazard functions have been considered in a number of earlier studies. The Taylor (1980) model was of course the first example of this kind and assumes that price changes take place at regular intervals. Guerrieri (2001, 2002) argues that the Taylor model actually fits empirical inflation dynamics better than the NKPC. Goodfriend and King (1997) show how it is possible to develop a theoretical model of time-dependent pricing with a general hazard function. The work of Dotsey, King and Wolman (1999) demonstrates that models of state-dependent pricing imply increasing hazard functions when there is base drift in the general price level. Some examples of these increasing hazard functions are studied by Wolman (1999) in the context of time-dependent pricing. Others have argued that a mixture of the Calvo and Taylor pricing yields a better model of inflation dynamics (Mash, 2004). It is also shown by Dotsey (2002) that econometric estimates of the "hybrid" NKPC of Galí and Gertler are biased towards detecting rule-of-thumb firms when prices are actually set according to the Taylor model. On the other hand, Fuhrer (1997) and Whelan (2007) take a different view and argue that more general time-dependent pricing models imply Phillips curves with negative coefficients on past inflation. This would suggest that having an increasing hazard function leads to a worse fit to the data than the NKPC. However, this claim is not supported by the results of this paper.

In interpreting these results, it is important to have a clear idea of what the models are trying to explain. Many studies judge success or failure in explaining inflation persistence using impulse response and autocorrelation functions for inflation. These are the metrics preponderantly used in statistical work on inflation persistence, as can be seen from Gadzinski and Orlandi (2004) and Altissimo, Bilke, Levin, Mathä and Mojon (2006). But the danger of this approach is aptly illustrated by Dittmar, Gavin and Kydland (2005), who argue that a model with entirely flexible prices can explain much of this reduced-form statistical evidence on inflation persistence. Hence this paper focuses on intrinsic inflation persistence, which is identified with lags of inflation in the Phillips curve having positive coefficients. As the expression for the Phillips curve derived here is much simpler than those typically found in previous studies, it can be shown more directly how an upward-sloping hazard function contributes to explaining intrinsic inflation persistence.

In addition to macroeconomic theorizing about the shape of the hazard function, there is now a wealth of microeconometric work that addresses this question. Unfortunately, the results of these studies are somewhat mixed. Götte, Minsch and Tyran (2005) and Cecchetti (1986) find strong support for an upward-sloping hazard function. Fougère, Le Bihan and Sevestre (2005) also find some support for increasing hazard functions for the majority of goods and services. On the other hand, work by Campbell and Eden (2005), Dias, Marques and Santos Silva (2005) and others find strong evidence in favour of mainly downward-sloping hazards. Other studies such as Baumgartner, Glatzer, Rumler and Stiglbauer (2005) suggest that the hazard function is decreasing but punctuated by spikes. Nakamura and Steinsson's (2007) estimated hazard function is largely flat but with a large spike after one year. These studies differ considerably in their econometric methodology, the range of goods and services they include, and the countries and time periods they cover. Some of these estimates of the hazard function slope may be biased downward as a result of not controlling adequately for heterogeneity (Álvarez, Burriel and Hernando, 2005).

Nonetheless, it is very important to be able to check the consistency of estimated hazard functions at the micro level with those derived from macroeconomic data, and estimates based on macro data are extremely rare in the literature. The estimation method proposed here is new, and the only other attempt based on macroeconomic data is Jadresic (1999). But because the estimation technique used by Jadresic is based on ordinary least squares, its validity rests on the very strong assumption of perfect foresight, rather than just on rational expectations as is needed in this paper.

The plan of the paper is as follows. The assumptions of the model are set out in section 2 and a simple expression for the implied Phillips curve is derived in section 3. This is then used to derive analytical results linking the shape of the hazard function to the extent of intrinsic inflation persistence. Section 4 describes the estimation procedure for the hazard function and presents estimates obtained using U.S. macroeconomic data. It goes on to assess how well the model can account for inflation dynamics and to compare the results with those obtained in the microeconometric literature. Finally, section 5 draws some conclusions.

2. The model

2.1 Firms' costs and demand

The economy contains a continuum of firms on the unit interval $\Omega \equiv [0, 1]$. Each firm produces a differentiated good that is an imperfect substitute for the products of other firms. The output of firm $\imath \in \Omega$ at time t is denoted by $Y_t(\imath)$. Firm $\imath$ can produce output $Y_t(\imath)$ at total real cost $C(Y_t(\imath); Y_t^*)$, given by

$$C(Y_t(\imath); Y_t^*) \equiv \frac{1}{1+\eta_{cy}} \frac{Y_t(\imath)^{1+\eta_{cy}}}{Y_t^{*\eta_{cy}}} \tag{1}$$

where the parameter $\eta_{cy} > 0$ is the elasticity of real marginal cost with respect to the firm's own output, and Y_t^* denotes the common Pareto-efficient level of output for all firms, which corresponds to the level of output where real marginal cost is equal to one. Efficient output Y_t^* depends on factors such as technology and household preferences, though it is not modelled here explicitly.[2] Firms take efficient output as exogenously given.

Firms' customers (households, government or other firms) allocate their spending between different goods to minimize the cost of buying some quantity of a basket of goods. Aggregate output Y_t is defined using a Dixit-Stiglitz aggregator:

$$Y_t \equiv \left(\int_\Omega Y_t(\imath)^{\frac{\varepsilon-1}{\varepsilon}} d\imath \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{2}$$

The parameter $\varepsilon > 1$ is the elasticity of substitution between different products. If $P_t(\imath)$ is the money price of firm $\imath$'s product then expenditure minimization by its customers implies that it faces the following demand function at time t,

$$Y_t(\imath) = \left(\frac{P_t(\imath)}{P_t} \right)^{-\varepsilon} Y_t \quad , \quad P_t \equiv \left(\int_\Omega P_t(\imath)^{1-\varepsilon} d\imath \right)^{\frac{1}{1-\varepsilon}} \tag{3}$$

where P_t is the corresponding price index for the basket of goods (2).

Using the demand function in (3) and the cost function in (1), let the level of real profits earned by firm $\imath$ at time t if its relative price is $\varrho_t(\imath) \equiv P_t(\imath)/P_t$ be denoted by $F(\varrho_t(\imath); Y_t, Y_t^*) = \varrho_t(\imath)^{1-\varepsilon} Y_t - C(\varrho_t(\imath)^{-\varepsilon} Y_t; Y_t^*)$. By substituting in the functional form from equation (1) and by defining the output gap $\mathcal{Y}_t \equiv Y_t/Y_t^*$ to be ratio of actual aggregate output to efficient output, real profits are given by:

$$F(\varrho_t(\imath); Y_t, Y_t^*) = \left\{ \varrho_t(\imath)^{1-\varepsilon} - \frac{1}{1+\eta_{cy}} \varrho_t(\imath)^{-\varepsilon(1+\eta_{cy})} \mathcal{Y}_t^{\eta_{cy}} \right\} Y_t \tag{4}$$

[2] See section 4.1 below for an example of how (1) can be derived.

2.2 Price stickiness

Instead of choosing relative prices directly, firms post prices in terms of money, and these money prices are not adjusted at every possible point in time. The frequency of price adjustment is modelled using the framework of time-dependent pricing, where a firm's probability of choosing a new price depends on the time elapsed since its previous price change.

Let $\mathcal{A}_t \subset \Omega$ denote the set of firms that post new prices at time t. The duration of price stickiness $\mathcal{D}_t(\imath) \equiv \min\left\{ i \geq 0 \,\middle|\, \imath \in \mathcal{A}_{t-i} \right\}$ for firm $\imath$ at time t is the time elapsed since its current price was posted. A particular model of time-dependent pricing is defined by a hazard function: a relationship between the probability of price adjustment and the duration of price stickiness. The hazard function is represented by a sequence of probabilities $\{\alpha_i\}_{i=1}^{\infty}$, with α_i denoting the probability of a firm posting a new price if its previous price change occurred i periods ago. The hazard function is formally defined by:

$$\alpha_i \equiv \mathbb{P}\left(\mathcal{A}_t \,\middle|\, \mathcal{D}_{t-1} = i - 1\right) \tag{5}$$

Every hazard function is associated with a corresponding survival function, a sequence $\{\varsigma_i\}_{i=0}^{\infty}$, where ς_i denotes the probability that a price posted at time t will still be in use at time $t + i$. There is a simple relationship between the hazard and survival functions:

$$\varsigma_i = \prod_{j=1}^{i}(1 - \alpha_j) \quad , \quad \varsigma_0 = 1 \tag{6}$$

Some weak restrictions need to be imposed on the hazard function:

Assumption 1 *The hazard function $\{\alpha_i\}_{i=1}^{\infty}$ is a sequence of well-defined probabilities $0 \leq \alpha_i \leq 1$, which satisfies the following restrictions:*
(i) There is some non-zero probability of price stickiness: $\alpha_1 < 1$
(ii) The probability of price adjustment is never exactly zero: $\alpha_i \geq \underline{\alpha} > 0$ for all $i = 1, 2, \ldots$.

Instead of specifying the entire hazard function $\{\alpha_i\}_{i=1}^{\infty}$ directly, this paper presents a new class of hazard functions where the shape is modelled parsimoniously using a small set of parameters. There is one parameter α to control the overall level of the hazard function and a set of n parameters $\{\varphi_i\}_{i=1}^{n}$ to control its slope. Greater flexibility in specifying the shape of the hazard function is obtained by increasing n. Using these parameters, the sequence of price-adjustment probabilities $\{\alpha_i\}_{i=1}^{\infty}$ is generated by the following recursion:[3]

$$\alpha_i = \alpha + \sum_{j=1}^{\min\{i-1,n\}} \varphi_j \left(\prod_{k=i-j}^{i-1}(1 - \alpha_k)\right)^{-1} \tag{7}$$

[3] Note that if the maximum duration of price stickiness $m \equiv \min\{\ i \mid \alpha_{i+1} = 1\ \}$ implied by (7) is finite, then the terms of the hazard function corresponding to prices older than m periods can be set to one without loss of generality.

Although the recursion (7) for the hazard function is non-linear, it is equivalent to a linear recursion for the corresponding survival function $\{\varsigma_i\}_{i=0}^{\infty}$:

$$\varsigma_i = (1-\alpha)\varsigma_{i-1} - \sum_{j=1}^{\min\{i-1,n\}} \varphi_j \varsigma_{i-1-j} \quad , \quad \varsigma_0 = 1 \tag{8}$$

The equivalence of (7) and (8) can be demonstrated using (6). It may be helpful to consider some examples of what (7) and (8) can be used to model.

Example 1 (All prices are equally sticky) The Calvo (1983) pricing model assumes that the probability of price adjustment is independent of the duration of price stickiness. In other words, the hazard function is constant with $\alpha_i = \alpha$ for some $0 < \alpha < 1$. Hence Calvo pricing is equivalent to the trivial case of a recursion with $n = 0$ in (7). The corresponding recursion for the survival function in (8) is first order implying that $\varsigma_i = (1-\alpha)^i$.

The simplest non-trivial example of a hazard function generated recursively using (7) is:

Example 2 (Newer prices are stickier than older prices) A hazard function that is upward sloping everywhere can be generated using a first-order recursion in (7). By setting $n = 1$, choosing $0 < \alpha < 1$ and a parameter φ between 0 and $(1-\alpha)^2/4$, a well-defined hazard function is obtained with $\alpha_i > \alpha_{i-1}$ for all i. The corresponding recursion for the survival function in (8) is second order with $\varsigma_i = (1-\alpha)\varsigma_{i-1} - \varphi\varsigma_{i-2}$.

A graphical illustration of the hazard and survival functions implied by Examples 1 and 2 is shown in Figure 1.

The formula (7) for $\{\alpha_i\}_{i=1}^{\infty}$ makes it clear that α_1 is always equal to α, so the parameter α always controls the initial level of the hazard function, the probability of adjusting a price posted one period ago. In Example 2, a positive value of the new parameter φ implies a positively sloped hazard function. The result given below shows that this principle extends to all the hazard functions generated by the recursion (7). Hence the parameters $\{\varphi_i\}_{i=1}^{n}$ are referred to as slope parameters, with positive values associated with positively sloped hazard functions and negative values with negatively sloped hazards.

Proposition 1 *Suppose that the hazard function $\{\alpha_i\}_{i=1}^{\infty}$ defined in (5) is generated by the recursion (7) using parameters α and $\{\varphi_i\}_{i=1}^{n}$. Then the slope of the hazard function $\Delta\alpha_{i+1}$ is connected to the signs of the parameters $\{\varphi_i\}_{i=1}^{n}$ as follows:*

(a) ***Flat hazard** :* $\varphi_i = 0$ *for all* $i = 1, \ldots, h \iff \alpha_{i+1} = \alpha_i = \alpha$ *for all* $i = 1, \ldots, h$

(b) ***Upward-sloping hazards** :*

i. $\varphi_i \geq 0$ *for all* $i = 1, \ldots, h \implies \alpha_{i+1} \geq \alpha_i$ *for all* $i = 1, \ldots, h$

ii. $\varphi_j > 0$ *for some* $j = 1, \ldots, h \impliedby \alpha_{i+1} > \alpha_i$ *for all* $i = 1, \ldots, h$

iii. $\varphi_h > 0 \implies \alpha_{j+1} > \alpha_j$ *for some* $j = 1, \ldots, h$

(c) ***Downward-sloping hazards** :*

i. $\varphi_i \leq 0$ *for all* $i = 1, \ldots, h \quad \Longleftarrow \quad \alpha_{i+1} \leq \alpha_i$ *for all* $i = 1, \ldots, h$

ii. $\varphi_i < 0$ *for all* $i = 1, \ldots, h \quad \Longrightarrow \quad \alpha_{j+1} < \alpha_j$ *for some* $j = 1, \ldots, h$

iii. $\varphi_j < 0$ *for some* $j = 1, \ldots, h \quad \Longleftarrow \quad \alpha_{h+1} < \alpha_h$

Proof. See appendix A.2. ■

The use of the recursion (7) to generate the hazard function raises two technical questions. First, whether every hazard function satisfying the weak restrictions in Assumption 1 can be represented by a recursion of the form (7). Secondly and conversely, whether every recursion (7) generates a hazard function satisfying Assumption 1. In brief, the respective answers are yes, approximately; and no, but restrictions on α and $\{\varphi_i\}_{i=1}^{n}$ can be found to check whether Assumption 1 is satisfied or not. These answers are justified formally by Propositions 2 and 3 below.

Proposition 2 *If a given hazard function* $\{\alpha_i\}_{i=1}^{\infty}$ *satisfies all the requirements of Assumption 1 then:*

(i) There exists a parameter α *and a sequence* $\{\varphi_i\}_{i=1}^{n}$ *of some length n (possibly infinite) such that these parameters exactly generate the original hazard function* $\{\alpha_i\}_{i=1}^{\infty}$ *using the recursion* (7).

(ii) If $n = \infty$ *then the sequence of parameters* $\{\varphi_i\}_{i=1}^{\infty}$ *is such that* $\lim_{i\to\infty} \varphi_i = 0$.

(iii) If $\{\alpha_i^{[h]}\}_{i=1}^{\infty}$ *is the hazard function generated by* (7) *but using only the first h terms of the sequence* $\{\varphi_i\}_{i=1}^{n}$, *then the first* $h + 1$ *terms of* $\{\alpha_i^{[h]}\}_{i=1}^{\infty}$ *agree exactly with those of* $\{\alpha_i\}_{i=1}^{\infty}$.

Proof. See appendix A.3. ■

Informally, Proposition 2 states that although high-order recursions are needed to represent every possible hazard function, the magnitude of the extra parameters required eventually becomes very small, so relatively low-order recursions can approximate a wide range of differently shaped hazard functions.

The second result concerns the restrictions on the parameters necessary and sufficient for the hazard function to satisfy Assumption 1. An illustrative result applying to first-order recursions ($n = 1$) is given below.[4]

Proposition 3 *Suppose the hazard function* $\{\alpha_i\}_{i=1}^{\infty}$ *is generated from parameters* α *and* φ *using* (7) *when* $n = 1$. *Then the resulting hazard function is well defined and satisfies Assumption 1 if and only if:*

$$0 < \alpha < 1 \quad , \quad -\frac{1}{4} < -\alpha(1-\alpha) \leq \varphi \leq \frac{1}{4}(1-\alpha)^2 < \frac{1}{4} \tag{9}$$

Proof. See appendix A.4. ■

[4] A general result can be derived that applies to all orders of recursion, though it is more complicated to use.

2.3 Profit-maximizing, forward-looking price setting

When firms anticipate that the prices they set are likely to be in use for several periods, it is necessary to balance maximizing profits today with profits in the future when selecting the best price. Suppose that at time t a firm posts a new money price, referred to as a reset price and denoted by R_t. If this price is still in use at time $\tau \geq t$ then the firm's relative price will be R_t/P_τ and it will earn profits $F(R_t/P_\tau; Y_\tau, Y_\tau^*)$ in real terms at that time, where the profit function is specified in equation (4). Future profits are discounted by financial markets using risk-free nominal interest rates.[5] Future profits also have to be discounted using the survival function $\{\varsigma_i\}_{i=0}^{\infty}$ because a new price might be posted before some of these profits are actually realized. The objective function of a firm choosing a reset price at time t is

$$\mathcal{F}_t \equiv \max_{R_t} \sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \mathbb{E}_t \left[\left(\prod_{s=t+1}^{\tau} \frac{\Pi_s}{\mathcal{I}_s} \right) F\left(\frac{R_t}{P_\tau}; Y_\tau, Y_\tau^* \right) \right] \tag{10}$$

where $\Pi_t \equiv P_t/P_{t-1}$ is the gross inflation rate between periods $t-1$ and t, $\mathcal{I}_t$ is the gross nominal interest rate also between $t-1$ and t, and $\mathbb{E}_t[\cdot]$ is the mathematical expectation operator conditional on all information available at time t. The first-order condition for the profit-maximizing reset price is obtained by differentiating (10) with respect to R_t,

$$\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \mathbb{E}_t \left[\left(\prod_{s=t+1}^{\tau} \frac{1}{\mathcal{I}_s} \right) F_\varrho\left(\frac{R_t}{P_\tau}; Y_\tau, Y_\tau^* \right) \right] = 0 \tag{11}$$

where $F_\varrho(\varrho_t(\imath); Y_t, Y_t^*)$ is the derivative of the profit function (4) with respect to the firm's own relative price:

$$F_\varrho(\varrho_t(\imath); Y_t, Y_t^*) = (1-\varepsilon)\left\{ 1 - \frac{\varepsilon}{\varepsilon - 1} \varrho_t(\imath)^{-(1+\varepsilon\eta_{cy})} \mathcal{Y}_t^{\eta_{cy}} \right\} \varrho_t(\imath)^{-\varepsilon} Y_t \tag{12}$$

Now let $x_t \equiv C_Y(Y_t, Y_t^*)$ denote the level of real marginal cost in a firm producing output equal to the economy-wide average Y_t. An expression for firm-specific real marginal cost $C_Y(Y_t(\imath); Y_t^*)$ is obtained from (1), which shows that x_t is an increasing function of the current output gap $\mathcal{Y}_t \equiv Y_t/Y_t^*$:

$$C_Y(Y_t(\imath); Y_t^*) = \left(\frac{Y_t(\imath)}{Y_t^*} \right)^{\eta_{cy}} \quad , \quad x_t = \mathcal{Y}_t^{\eta_{cy}} \tag{13}$$

The profit-maximizing reset price R_t is obtained by combining equations (11)–(13),

$$R_t = P_t \left(\frac{\frac{\varepsilon}{\varepsilon-1} \sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \mathbb{E}_t \left[\left(\prod_{s=t+1}^{\tau} \left(\frac{G_s \Pi_s^{\varepsilon}}{\mathcal{I}_s} \right) \Pi_s^{(1+\varepsilon\eta_{cy})} \right) x_\tau \right]}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \mathbb{E}_t \left[\prod_{s=t+1}^{\tau} \frac{G_s \Pi_s^{\varepsilon}}{\mathcal{I}_s} \right]} \right)^{\frac{1}{1+\varepsilon\eta_{cy}}} \tag{14}$$

[5] Discounting profits using a more general stochastic discount factor would not change the results presented here.

where $G_t \equiv Y_t/Y_{t-1}$ denotes the gross growth rate of aggregate real output. The optimal reset price is a weighted average of current and future real marginal costs and inflation rates. Note that all firms choosing a reset price at the same time have an incentive to pick the same value of R_t appearing in (14).

2.4 Aggregation

Denote the distribution of the duration of price stickiness at time t using the sequence $\{\theta_{it}\}_{i=0}^{\infty}$, where $\theta_{it} \equiv \mathbb{P}(\mathcal{D}_t = i)$ is the proportion of firms using a price set i periods ago. The definition of the hazard function implies this distribution evolves over time according to:

$$\theta_{0t} = \sum_{i=1}^{\infty} \alpha_i \theta_{i-1,t-1} \quad , \quad \theta_{it} = (1-\alpha_i)\theta_{i-1,t-1} \qquad i = 1, 2, \ldots \tag{15}$$

If the hazard function satisfies Assumption 1 then the scope for time-variation in the distribution $\{\theta_{it}\}_{i=0}^{\infty}$ is transitory: there is a unique stationary distribution to which the economy must converge.

Proposition 4 *Suppose that the hazard function $\{\alpha_i\}_{i=1}^{\infty}$ satisfies all the requirements of Assumption 1 and that the evolution over time of the distribution of the duration of price stickiness is given by* (15).

(i) There exists a unique stationary distribution $\{\theta_i\}_{i=0}^{\infty}$ to which the economy converges from any starting point.

(ii) Now suppose the hazard function is generated by the recursion in (7) *and assume that the economy has converged to the unique stationary distribution. Then the distribution of the duration of price stickiness is proportional to the survival function, and the unconditional probability of price adjustment $\alpha^e \equiv \sum_{i=1}^{\infty} \alpha_i \theta_{i-1}$ and the unconditional expected duration of price stickiness $\mathcal{D}^e \equiv \sum_{i=1}^{\infty} i\theta_{i-1}$ are given by:*

$$\theta_i = \left(\alpha + \sum_{j=1}^{n} \varphi_j\right) \varsigma_i \quad , \quad \alpha^e = \alpha + \sum_{i=1}^{n} \varphi_i \quad , \quad \mathcal{D}^e = \frac{1 - \sum_{i=1}^{n} i\varphi_i}{\alpha + \sum_{i=1}^{n} \varphi_i} \tag{16}$$

Proof. See appendix A.5. ■

In what follows, the economy is assumed to have converged to the unique stationary distribution $\{\theta_i\}_{i=0}^{\infty}$, so $\mathbb{P}(\mathcal{D}_t = i) = \theta_i$ for all t. The price level P_t defined in (3) can then be written in terms of a time-invariant weighted average of past reset prices:

$$P_t = \left(\sum_{i=0}^{\infty} \theta_i R_{t-i}^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \tag{17}$$

3. Inflation dynamics

For given stochastic processes for real marginal cost $\{x_t\}$, the growth rate of real output $\{G_t\}$ and the nominal interest rate $\{\mathcal{I}_t\}$, equations (14) and (17) determine the inflation rate $\{\Pi_t\}$. But as it is not generally possible to solve this system of non-linear equations analytically, the following section shows instead how a log-linear approximation to the solution can be found.

3.1 Steady state and log linearization

First note that equations (14) and (17) are homogeneous of degree zero in all prices expressed in money terms. These equations can then be recast in terms of the gross inflation rate $\Pi_t \equiv P_t/P_{t-1}$ and the relative reset price $r_t \equiv R_t/P_t$ as follows:

$$r_t = \left(\frac{\frac{\varepsilon}{\varepsilon-1} \sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \mathbb{E}_t \left[\left(\prod_{s=t+1}^{\tau} \frac{G_s \Pi_s^{\varepsilon+(1+\varepsilon\eta_{cy})}}{\mathcal{I}_s} \right) x_\tau \right]}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \mathbb{E}_t \left[\prod_{s=t+1}^{\tau} \frac{G_s \Pi_s^{\varepsilon}}{\mathcal{I}_s} \right]} \right)^{\frac{1}{1+\varepsilon\eta_{cy}}} \quad , \quad 1 = \sum_{i=0}^{\infty} \theta_i r_{t-i}^{1-\varepsilon} \left(\prod_{j=0}^{i-1} \Pi_{t-j}^{\varepsilon-1} \right) \tag{18}$$

It is straightforward to check that given a trend inflation rate ($\Pi_t = \bar{\Pi}$), a trend rate of output growth ($G_t = \bar{G}$) and a steady-state nominal interest rate $\mathcal{I}_t = \bar{\mathcal{I}}$, (18) implies a well-defined steady state for the relative reset price ($r_t = \bar{r}$) and real marginal cost ($x_t = \bar{x}$). For simplicity, the model is log-linearized around a steady state with zero inflation ($\bar{\Pi} = 1$) and zero real output growth ($\bar{G} = 1$), which leads to a steady state with $\bar{r} = 1$ and $0 < \bar{x} < 1$.[6] The steady-state real interest rate is assumed positive and is represented by $\bar{\mathcal{I}}/\bar{\Pi} = \beta^{-1}$, where β is a discount factor satisfying $0 < \beta < 1$.

In what follows, log deviations of variables from their steady-state values are denoted by sans serif letters. For variables that are indeterminate in the steady state, the sans serif letter simply denotes the logarithm. The equations in (18) can be log-linearized around the steady state defined above,

$$\mathsf{R}_t = \sum_{i=0}^{\infty} \left(\frac{\beta^i \varsigma_i}{\sum_{j=0}^{\infty} \beta^j \varsigma_j} \right) \mathbb{E}_t \left[\mathsf{P}_{t+i} + \eta_{cx} \mathsf{x}_{t+i} \right] \quad , \quad \mathsf{P}_t = \sum_{i=0}^{\infty} \theta_i \mathsf{R}_{t-i} \tag{19}$$

where the parameter $\eta_{cx} \equiv 1/(1+\varepsilon\eta_{cy})$ represents the sensitivity of an individual firm's marginal cost to average real marginal cost x_t when it keeps its price constant.

3.2 The Phillips curve

The conventional approach to deriving the Phillips curve implied by a model of time-dependent pricing is to combine the two equations in (19), eliminate the reset price R_t, and recast the equation in terms of inflation $\pi_t \equiv \mathsf{P}_t - \mathsf{P}_{t-1}$ and real marginal cost x_t. This has a number

[6] This steady state is chosen for simplicity in many New Keynesian models. It is not difficult to extend the results in this paper to cases where $\bar{\Pi} \neq 1$ or $\bar{G} \neq 1$.

of drawbacks. First, the resulting equation has a complicated autoregressive distributed lag structure in inflation, real marginal cost, and conditional expectations of both variables subject as many different information sets as the maximum duration of price stickiness. This makes intrinsic inflation persistence very difficult to characterize, as any inflation persistence implied by this Phillips curve could derive from either the lags of inflation, the lags of real marginal cost, the lagged expectations of both, as well as from persistence in real marginal cost itself. Furthermore, because of the presence of expectations of the same variable conditional on many different information sets, and the need for as many lags and leads of each variable as the maximum duration of price stickiness, the Phillips curve equation is extremely difficult to estimate using the limited-information techniques such as GMM that have proved so popular for the New Keynesian Phillips curve.

This paper considers an simpler alternative expression for the Phillips curve which circumvents these problems of interpretation and estimation. The key to obtaining a simple Phillips curve is to exploit the recursive parameterization of the hazard function $\{\alpha_i\}_{i=1}^{\infty}$ in (7), which Proposition 2 shows can be used quite generally. With this parameterization, equation (8) characterizes the survival function $\{\varsigma_i\}_{i=0}^{\infty}$, which implies that the equation for the profit-maximizing reset price in (19) can be replaced by the following recursive equivalent:

$$\mathsf{R}_t = \beta(1-\alpha)\mathbb{E}_t\mathsf{R}_{t+1} - \sum_{i=1}^{n}\beta^{i+1}\varphi_i\mathbb{E}_t\mathsf{R}_{t+1+i} + \left(1-\beta(1-\alpha)+\sum_{i=1}^{n}\beta^{i+1}\varphi_i\right)(\mathsf{P}_t+\eta_{cx}\mathsf{x}_t) \qquad (20)$$

To apply the same approach to the price level equation, note that the result in equation (16) of Proposition 4 shows that the recursive parameterization of the hazard function implies a recursion for the stationary distribution of the duration of price stickiness $\{\theta_i\}_{i=0}^{\infty}$:

$$\theta_i = (1-\alpha)\theta_{i-1} - \sum_{j=1}^{\min\{i-1,n\}}\varphi_j\theta_{i-j-1} \quad , \quad \theta_0 = \alpha + \sum_{j=1}^{n}\varphi_j \qquad (21)$$

Hence the price level equation in (19) can be replaced by the following recursive equivalent:[7]

$$\mathsf{P}_t = (1-\alpha)\mathsf{P}_{t-1} - \sum_{i=1}^{n}\varphi_i\mathsf{P}_{t-1-i} + \left(\alpha+\sum_{i=1}^{n}\varphi_i\right)\mathsf{R}_t \qquad (22)$$

Solving equations (20) and (22) instead of those in (19) yields a much simpler expression for the Phillips curve.

Theorem 1 *Suppose that the hazard function $\{\alpha_i\}_{i=1}^{\infty}$ satisfies Assumption 1 and is generated by the n-th order recursion* (7) *with parameters α and $\{\varphi_i\}_{i=1}^{n}$. If the profit-maximizing reset price R_t is given by equation* (20) *and the price level P_t by* (22)*, and $0 < \beta < 1$ and $\eta_{cx} > 0$, then the Phillips curve relationship between inflation $\pi_t = \mathsf{P}_t - \mathsf{P}_{t-1}$ and real marginal cost x_t*

[7] The translation of the equations in (19) into recursive versions (20) and (22) is essentially equivalent to finding autoregressive representations of invertible moving-average processes.

is:

$$\pi_t = \sum_{i=1}^{n} \psi_i \pi_{t-i} + \sum_{i=1}^{n+1} \delta_i \mathbb{E}_t \pi_{t+i} + \kappa_x \mathsf{x}_t \tag{23}$$

Inflation π_t depends on current real marginal cost x_t, n lags of inflation and $n+1$ values of expected future inflation, where n is the order of the recursion that generates the hazard function in (7)*. Note that real marginal cost can be replaced by the output gap using $\mathsf{x}_t = \eta_{cy}\mathsf{y}_t$, a log-linearized version of* (13)*. The coefficients of lagged inflation $\{\psi_i\}_{i=1}^{n}$ and future inflation $\{\delta_i\}_{i=1}^{n+1}$ depend on the parameters α, $\{\varphi_i\}_{i=1}^{n}$ and β. The coefficient κ_x on real marginal cost depends on η_{cx} in addition to these.*

The signs of the coefficients on lagged inflation are determined by the hazard function slope parameters $\{\varphi_i\}_{i=1}^{n}$:

(a) ***Flat hazard*** *: $\varphi_i = 0$ for all $i = 1, \ldots, n \iff \psi_i = 0$ for all $i = 1, \ldots, n$*

(b) ***Upward-sloping hazards*** *:*

- *i. $\varphi_i > 0$ for all $i = 1, \ldots, n \implies \psi_i > 0$ for all $i = 1, \ldots, n$*
- *ii. $\varphi_i > 0$ for some i and $\varphi_j = 0$ for all $j \neq i \implies \psi_j > 0$ for all $j = 1, \ldots, i$*
- *iii. $\varphi_j > 0$ for some $j = 1, \ldots, n \impliedby \psi_i > 0$ for some $i = 1, \ldots, n$*

(c) ***Downward-sloping hazards*** *:*

- *i. $\varphi_i < 0$ for all $i = 1, \ldots, n \implies \psi_i < 0$ for all $i = 1, \ldots, n$*
- *ii. $\varphi_i < 0$ for some i and $\varphi_j = 0$ for all $j \neq i \implies \psi_j < 0$ for all $j = 1, \ldots, i$*
- *iii. $\varphi_j < 0$ for some $j = 1, \ldots, n \impliedby \psi_i < 0$ for some $i = 1, \ldots, n$*

There are also $n+1$ restrictions linking the coefficients of past inflation $\{\psi_i\}_{i=1}^{n}$ and the discount factor β to the coefficients of future inflation $\{\delta_i\}_{i=1}^{n+1}$. These hold for all hazard functions:

$$\delta_1 = \beta + (1-\beta)\sum_{j=1}^{n}\beta^j\psi_j \quad , \quad \delta_i = -\left(\beta^i\psi_{i-1} - (1-\beta)\sum_{j=i}^{n}\beta^j\psi_j\right) \qquad i = 2, \ldots, n+1 \tag{24}$$

Proof. See appendix A.6. ■

The key insight here is that the presence of lags of inflation in (23) is perfectly consistent with purely forward-looking firms whenever the hazard function is not flat, that is, whenever a non-trivial recursion (7) is used with $n \geq 1$. And more importantly, these lags of inflation have positive coefficients when the hazard function is upward sloping.

The intuition for these findings can be understood by considering the effects of a shock that initially increases inflation. Because price-adjustment times are staggered, only a subset of firm increases their prices to begin with. Since all firms that subsequently change price at the same time choose a common reset price, those firms that did not change price at the first onset of the shock have further to catch up than those that have already responded to the shock. So if a larger proportion of subsequent price changes come from those firms that made no price change initially then the rate of inflation will be higher in the periods after the arrival of the shock. This is precisely what happens when the hazard function is upward sloping: newly set prices are less likely to be changed than those that have been left fixed for a long time. The extra inflation persistence created in this case relative to a flat hazard function is captured by the presence of lagged inflation rates with positive coefficients.

One simple illustration of the results of Theorem 1 is given by a comparison of the Phillips curves implied by Examples 1 and 2. In Example 1, the hazard function is completely flat, which is the key assumption of the Calvo (1983) model of sticky prices. Such a hazard function can be generated by the trivial case of a recursion with $n = 0$ in (7). In this special case, Theorem 1 simply reproduces the well-known result that the Calvo pricing model implies the New Keynesian Phillips curve,

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \left(\frac{\alpha(1 - \beta(1 - \alpha))\eta_{cx}}{1 - \alpha} \right) \mathsf{x}_t \tag{25}$$

where α is the constant probability of price adjustment.[8] The NKPC states that inflation depends only on the current level of real marginal cost and expected inflation one period in the future, and it has attracted criticism because it lacks any role for past inflation. A popular alternative empirical specification, which nests the NKPC, is the so-called hybrid New Keynesian Phillips curve (HNKPC):

$$\pi_t = \mathfrak{b}_p \pi_{t-1} + \mathfrak{b}_f \mathbb{E}_t \pi_{t+1} + \mathfrak{b}_x \mathsf{x}_t \tag{26}$$

This alternative Phillips curve postulates that past inflation is an explicit determinant of current inflation, but it has proved more difficult to find a readily acceptable theoretical foundation for $\mathfrak{b}_p > 0$ in (26).

Now consider the upward-sloping hazard function of Example 2 again. It is generated by a first-order hazard function recursion using two parameters α and φ, with the former controlling the level of the hazard function and the latter its slope. According to Theorem 1, the Phillips curve in this case takes the form:

$$\pi_t = \psi \pi_{t-1} + \beta(1 + (1 - \beta)\psi)\mathbb{E}_t \pi_{t+1} - \beta^2 \psi \mathbb{E}_t \pi_{t+2} + \kappa_x \mathsf{x}_t \tag{27}$$

Current inflation now depends directly on past inflation, along with real marginal cost and expected inflation one and two periods in the future. Apart from the second future inflation term, (27) has the same form as the hybrid New Keynesian Phillips curve (26), but unlike the latter, it has a clear theoretical foundation.

The coefficient ψ determines the weight attached to past inflation relative to expected future inflation in influencing current inflation. When β is close to one, the weights on past and future inflation are approximately ψ and $1 - \psi$ respectively. A positive value of ψ means that current inflation depends positively on lagged inflation and that the weight on future inflation is reduced, though remaining positive if $\psi < 1$. While expected inflation two periods in the future

[8] See Woodford (2003) for a detailed derivation of the NKPC and further discussion.

then has a negative coefficient, it should not be interpreted as implying that higher expected inflation two periods ahead lowers inflation today. This is because the Phillips curve (27) at time $t+1$ implies that a rise in inflation in period $t+2$ creates a similar amount of inflationary pressure in period $t+1$. The sum of the coefficients on both future inflation terms is positive if $\psi < 1$, so in this case the expected inflation in period $t+2$ would still raise inflation today, albeit by less than when $\psi = 0$. Therefore the negative coefficient on expected inflation two periods in the future should be interpreted only as a reduction in the overall weight attached to future inflation.

The coefficient ψ of lagged inflation and the coefficient κ_x of real marginal cost are obtained from the following functions of the parameters α, φ, β and η_{cx}:

$$\psi = \frac{\varphi}{(1-\alpha) - \varphi(1-\beta(1-\alpha))} \quad , \quad \kappa_x = \frac{(\alpha+\varphi)(1-\beta(1-\alpha)+\beta^2\varphi)\eta_{cx}}{(1-\alpha) - \varphi(1-\beta(1-\alpha))} \tag{28}$$

For the underlying hazard function to be well defined, the inequalities in (9) involving the parameters α and φ must be satisfied. These guarantee that κ_x is always positive, and that the sign of ψ depends only on the sign of φ. As Proposition 1 shows that $\varphi > 0$ implies a hazard function that is positively sloped everywhere, it is seen that this type of hazard function leads to a Phillips curve in which lagged inflation has a positive coefficient $\psi > 0$. The magnitude of this coefficient is increasing in the hazard-function slope parameter φ.

The principle that a positively sloped hazard function is associated with positive coefficients of lagged inflation generalizes to richer models with more parameters. This is because the same set of parameters $\{\varphi_i\}_{i=1}^n$ controls both the slope of the hazard function according to Proposition 1, and the signs of the coefficients of lagged inflation in the Phillips curve (23) according to Theorem 1. This is the key theoretical result contained in this paper, which is stated below formally:

Corollary 1 *Suppose firms maximize profits* (10) *when they set prices and that the hazard function for price adjustment* $\{\alpha_i\}_{i=1}^\infty$ *satisfies Assumption 1.*

(i) There exists a class of hazard functions that are everywhere upward sloping and which imply that all the coefficients of past inflation in the Phillips curve (23) *are positive.*

(ii) If one or more of the coefficients of lagged inflation in the Phillips curve (23) *is positive then the hazard function must have one or more upward-sloping sections.*

Proof. These claims follow immediately from Propositions 1 and 2 together with Theorem 1. ■

Thus by constructing a hazard function using (7) with all the φ_i parameters positive, it is possible to explain any number of positive coefficients of lagged inflation. In fact, as Proposition 1 and Theorem 1 show, an appropriate choice of signs for the parameters $\{\varphi_i\}_{i=1}^n$ can generate essentially any pattern of signs for the sequence of coefficients of past inflation. Therefore, on its own, the hypothesis of time-dependent pricing with forward-looking, profit-maximizing firms has no implications for the signs of the coefficients of past inflation. But when combined with a hypothesis about the shape of the hazard function, there are clear implications for the signs of these coefficients. Corollary 1 shows that there are hazard functions which are upward sloping everywhere that imply Phillips curves in which every coefficient of lagged inflation is positive. This sufficiency result is complemented by a corresponding necessity result. If at least one of the coefficients of lagged inflation is positive and firms are forward-looking profit maximizers then the hazard function must be positively sloped somewhere. It follows immediately that if the hazard function were everywhere downward sloping then all the coefficients on lagged inflation would be unambiguously negative.

4. Estimating the hazard function

The analysis in section 3 demonstrates that the shape of the hazard function is systematically related to inflation dynamics. By exploiting this insight it is possible to devise a method for estimating the hazard function that requires only macroeconomic data and simple econometric techniques. No individual price observations are needed. This method is used to answer the question of whether a hazard function model can be found that is quantitatively as well as qualitatively consistent with the behaviour of inflation. The results are also compared with those derived from the more conventional microeconometric approach.

4.1 Estimation method and specification issues

An observable proxy for real marginal cost

It is first necessary to find some observable proxy for the driving variable x_t in the Phillips curve (23), that is, for the level of real marginal cost in the average firm. One solution is to replace it with the output gap y_t, which then in practice could be equated with the deviation of aggregate output from some trend. This approach is eschewed here for a number of reasons. First, the link between average real marginal cost and the output gap derived in section 3.1 may change in the presence of features such as sticky wages or risk-sharing employment contracts from which this paper has abstracted. Second, there are many different detrending procedures for aggregate output and thus a range of "output gap" measures to choose from. It is difficult to know which statistical detrending procedure, if any, delivers a measure consistent with the theoretical concept of the output gap required by the model.

An alternative approach that has become popular in work on the New Keynesian Phillips curve is to use (real) unit labour costs in place of real marginal cost (Galí and Gertler, 1999; Sbordone, 2002). Unit labour costs have the advantage of being readily measurable without the need for detrending. Justifying the substitution of unit labour costs for real marginal cost requires a few more assumptions beyond those introduced in section 2.1. Assume each firm $\imath \in \Omega$ in the economy faces a Cobb-Douglas production function $Y_t(\imath) = A_t H_t(\imath)^{\eta_{yh}}$. Firm $\imath$

produces output $Y_t(\imath)$ by using $H_t(\imath)$ hours of a homogeneous labour input. The term A_t captures exogenous technological progress, and the parameter η_{yh} measures the elasticity of output with respect to hours and is assumed to satisfy $0 < \eta_{yh} \leq 1$. Firms can hire as many hours of labour as they want at real wage w_t.

It is not difficult to show that the Cobb-Douglas production function and the assumption that firms are wage takers implies that the total real cost function $C\left(Y_t(\imath); Y_t^*\right)$ takes the form given in equation (1), with parameter $\eta_{cy} = (1 - \eta_{yh})/\eta_{yh}$ measuring the elasticity of firms' real marginal cost with respect to their own output. If H_t is the total number of hours supplied by all workers then the combination of the labour market equilibrium condition, the individual Cobb-Douglas production functions for each firm and the demand curves in (3) implies an aggregate production function $Y_t = A_t(H_t/\varDelta_t)^{\eta_{yh}}$, where $\varDelta_t$ is an index of relative-price dispersion.

Average real marginal cost is defined by $x_t \equiv C_Y\left(Y_t, Y_t^*\right)$, and given the aggregate production function and (13) it follows that $x_t = (w_t H_t)/(\eta_{yh} Y_t \varDelta_t)$. Hence if $\mathrm{s}_t \equiv w_t H_t / Y_t$ is the labour share of income then $x_t = \mathrm{s}_t/(\eta_{yh}\varDelta_t)$. Since the first-order terms of a Taylor expansion of $\varDelta_t$ around the steady state defined in section 3.1 are all zero, the log deviation x_t of real marginal cost from its steady-state value is equal to the log deviation s_t of the labour share, ignoring second- and higher-order terms. In the data, unit labour costs are defined as labour compensation divided by output, so when expressed in real terms this measure is equivalent to the labour share of income, and hence to real marginal cost, under the assumptions made in this section.

Identification

The approach to identifying the hazard function $\{\alpha_i\}_{i=1}^{\infty}$ using macroeconomic data relies on the connection between the coefficients $\{\psi_i\}_{i=1}^{n}$, $\{\delta_i\}_{i=1}^{n+1}$ and κ_x appearing in the Phillips curve (23) and the hazard function parameters α and $\{\varphi_i\}_{i=1}^{n}$.

Suppose first that the Phillips curve coefficients are identified. There are $2(n + 1)$ of these coefficients, and $n + 3$ underlying parameters to be identified, namely α, $\{\varphi_i\}_{i=1}^{n}$, β and η_{cx}. A result of Theorem 1 is that each inflation coefficient ψ_i or δ_i is a function of α, $\{\varphi_i\}_{i=1}^{n}$ and the discount factor β. But the mapping between the coefficients and the parameters is non-linear, and it turns out that if any subset of size n of $\{\psi_i\}_{i=1}^{n}$ and $\{\delta_i\}_{i=1}^{n+1}$ is known then the remaining $n+1$ coefficients can be inferred given the value of β. So in total, the Phillips curve coefficients provide only $n + 2$ independent pieces of information about the parameters and thus an extra restriction is needed.

This extra information is provided by a calibration of the elasticities ε and η_{cy}, which pins down $\eta_{cx} = 1/(1 + \varepsilon\eta_{cy})$. The equations in (18) imply $\varepsilon/(\varepsilon - 1)\bar{x} = 1$, where $\bar{x}$ is steady-state real marginal cost. It follows that the model generates an average markup on marginal cost of $1/(\varepsilon - 1)$. By combining this result with the analysis from the previous sub-section, the model is seen to imply an average labour share of income of $(\varepsilon - 1)/(\varepsilon(1 + \eta_{cy}))$. The average markup is set to 20%, which implies a price elasticity of demand $\varepsilon = 6$. With the average labour share equal to 67%, an elasticity of real marginal cost with respect to output of $\eta_{cy} = 0.25$ is required. It follows that $\eta_{cx} = 0.4$. As a robustness check, the special case of $\eta_{cx} = 1$ is also considered. This was originally the only case considered in Galí and Gertler's (1999) work on the New Keynesian Phillips curve. The subsequent study by Galí, Gertler and López-Salido (2001) considers both $\eta_{cx} = 1$ and $0 < \eta_{cx} < 1$. From the results derived earlier, it is apparent that $\eta_{cx} = 1$ requires $\eta_{cy} = 0$, which is equivalent to no diminishing returns to labour in the short run. This seems implausible, so $\eta_{cx} = 0.4$ is the preferred choice here.

Even if the hazard function parameters α and $\{\varphi_i\}_{i=1}^n$ are identified, there is no guarantee that their estimated values will automatically imply a well-defined hazard function. Theorem 1 reveals that a hazard function satisfying Assumption 1 generated by (7) using α and $\{\varphi_i\}_{i=1}^n$ necessarily implies a Phillips curve of the form (23) for some coefficients $\{\psi_i\}_{i=1}^n$, $\{\delta_i\}_{i=1}^{n+1}$ and κ_x. But the converse is not true for the hazard function parameters recovered from any arbitrary set of Phillips curve coefficients. However, once the parameters are estimated, it is possible to test whether the implied hazard function is well defined or not.

The foregoing discussion assumes that the coefficients in the Phillips curve (23) are themselves identified. Suppose the Phillips curve equation (23) holds with an error term $\nu_t \sim \mathcal{IID}(0, \sigma_\nu^2)$, and that real marginal cost x_t is replaced by the observable labour share of income s_t (real unit labour cost) as explained earlier. If the expected future inflation rates are replaced by their realized values then the Phillips curve equation becomes $\pi_t = \sum_{i=1}^n \psi_i \pi_{t-i} + \sum_{i=1}^{n+1} \delta_i \pi_{t+i} + \kappa_x \mathsf{s}_t + \upsilon_t$, where $\upsilon_t \equiv \nu_t - \sum_{i=1}^{n+1} \delta_i \mathsf{e}_{t+i}^i$ is a composite error term that depends on the i-step ahead prediction errors $\mathsf{e}_t^i \equiv \pi_t - \mathbb{E}_{t-i}\pi_t$.

At time t, the variables $\{ \mathsf{s}_t, \pi_t, \ldots, \pi_{t+n+1} \}$ are endogenous and non-predetermined, so instruments are required to achieve identification of all the coefficients. Let $\mathbf{z}_{t-1}$ be a $q \times 1$ vector of observable variables that are known by firms at time $t-1$. If firms do not make predictable errors when forecasting inflation then υ_t should be uncorrelated with $\mathbf{z}_{t-1}$, implying the following moment conditions involving the coefficients $\{\psi_i\}_{i=1}^n$, $\{\delta_i\}_{i=1}^{n+1}$ and κ_x:

$$\mathbb{E}\left[\left(\pi_t - \sum_{i=1}^{n} \psi_i \pi_{t-i} - \sum_{i=1}^{n+1} \delta_i \pi_{t+i} - \kappa_x \mathsf{s}_t\right) \mathbf{z}_{t-1}\right] = \mathbf{0} \tag{29}$$

Identification of the hazard function parameters α and $\{\varphi_i\}_{i=1}^n$ and the discount factor β therefore requires there to be at least $n+2$ variables in $\mathbf{z}_{t-1}$ that have predictive power for the current and future endogenous variables appearing in (23). However, since the underlying theory itself implies that inflation is given by (23), n lags of inflation can always be included in $\mathbf{z}_{t-1}$. This leaves only two other variables to be found with the necessary predictive power.

Econometric technique

A limited-information approach is employed here to estimating the hazard function using a single Phillips curve equation. In particular, a generalized method of moments (GMM) estimator is applied using the moment conditions in (29).[9] The methodology mirrors that used by Galí and Gertler (1999) to estimate the New Keynesian Phillips curve.[10]

The hazard function parameters α and $\{\varphi_i\}_{i=1}^{n}$ are estimated using the moment conditions in (29) and the link with the Phillips curve coefficients $\{\psi_i\}_{i=1}^{n}$, $\{\delta_i\}_{i=1}^{n+1}$ and κ_x provided by Theorem 1. This means that the coefficients appearing in the moment conditions (29) are actually non-linear functions of the estimated parameters, and so the issue of the normalization of the moment conditions must be addressed. In small samples, the choice of normalization can affect the results. Rather than taking the moment conditions as they are in (29), these conditions are multiplied by a function of the parameters that ensures the resulting Phillips curve coefficients are bounded whenever the set of parameters α and $\{\varphi_i\}_{i=1}^{n}$ is bounded.[11] The alternative normalization that leaves (29) unchanged, which imposes a coefficient of one on current inflation, is used as a robustness check. An equivalent pair of normalizations is also considered by Galí and Gertler (1999), who claim that the bounded normalization is shown by simulation studies to have better small-sample properties than the normalization with a coefficient of one on current inflation. So the bounded normalization is the preferred specification here and is denoted by $\mathcal{N}(1)$. The alternative normalization on current inflation is denoted by $\mathcal{N}(2)$.

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.[12]

Data

Quarterly U.S. data from 1960:Q1 to 2003:Q4 are used.[13] Inflation is measured by the annualized percentage change in the GDP deflator between consecutive quarters. Real unit labour costs are given by unit labour costs in the business sector divided by the GDP deflator, and expressed as a percentage deviation from their average value. 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 unit labour costs themselves: the spread between ten-year Treasury Bond

[9]The estimation is performed using Cliff's (2003) GMM package for MATLAB.

[10]The alternative of full-information maximum likelihood estimation is not pursued here since it would require a complete model of the data-generating process, and would be less robust than GMM if this were misspecified. For maximum likelihood estimation of the NKPC, see Lindé (2005) and Kurmann (2004).

[11]The requisite factor is the expression in the denominators of the coefficients $\{\psi_i\}_{i=1}^{n}$, $\{\delta_i\}_{i=1}^{n+1}$ and κ_x, as given in the block of equations (A.4) from the proof of Theorem 1.

[12]For more details on different GMM estimation methods, see Mátyás (1999).

[13]The source of the data is the Federal Reserve Economic Data (FRED) database, which is available online at `research.stlouisfed.org/fred2`.

and three-month Treasury Bill yields, quadratically detrended log real GDP, the rate of wage inflation (annualized percentage change in compensation per hour in the business sector), and the rate of commodity-price inflation as measured by the percentage change between consecutive quarters of a futures-price index. These are very similar to the instruments used in the original Galí and Gertler (1999) study of the NKPC. Based on their statistical significance in a predictive regression for future inflation, six lags of inflation and commodity-price inflation are selected as instruments, together with two lags of each of the other variables.

4.2 Estimation results

To establish a benchmark with which later results can be compared, the hazard function is first estimated when it is constrained to be flat, as in Example 1. This is the Calvo pricing model underlying the standard New Keynesian Phillips curve. It is obtained by imposing $n = 0$ in the hazard function recursion (7). There are just two parameters to estimate: the constant probability of price adjustment α and the discount factor β. The parameter η_{cx} is calibrated as discussed in section 4.1. Estimates are presented in Table 1 for all pairings of the calibrated values of η_{cx} and the normalizations of the moment conditions detailed in section 4.1. The preferred specification is $\eta_{cx} = 0.4$ and $\mathcal{N}(1)$.

The constant probability of price adjustment is found to be 0.405 per quarter under the preferred specification $\eta_{cx} = 0.4$ and $\mathcal{N}(1)$. This is quite high, though by no means inconsistent with the micro-level evidence on price adjustment. As (25) shows, the coefficient of future inflation is determined entirely by the discount factor β. The estimates of β are not significantly different from one in any specification. The coefficient of unit labour costs s_t (the labour share) is positive and significant at the 5% level when the preferred normalization $\mathcal{N}(1)$ is used. Notice that the standard errors tend to be larger when normalization $\mathcal{N}(2)$ is used. As the hazard function is constrained to be flat, the estimates of the expected probability of price adjustment α^e are identical to the parameter α itself. The expected duration of price stickiness $\mathcal{D}^e$ is calculated using (16) and is found to be 2.469 quarters under the preferred specification. Under the alternative specifications, the expected duration is estimated to be noticeably longer. Finally, none of the J-statistics reports a rejection of the over-identifying moment conditions.[14]

The estimated hazard and survival functions for the Calvo model are plotted in Figure 2. These are derived from the estimated parameters under the preferred specification in Table 1. The hazard function is of course necessarily flat and the survival function decays at a constant geometric rate, as was first seen for Example 1 in Figure 1. The thick and thin bars in Figure 2 represent one-standard-deviation and two-standard-deviations bands around the point estimates.

The aim is now to use macroeconomic data to estimate the shape of the hazard function, imposing as few a priori restrictions as possible on what shapes are admissible. A first step

[14]The J-statistic is the Hansen test of over-identifying restrictions derived from surplus moment conditions. See Mátyás (1999) for further details.

towards this goal is taken by estimating a hazard function with both a level parameter α and one slope parameter φ_1. This is done by considering hazard functions in the class generated by first-order recursions, as illustrated by Example 2. This allows monotonically increasing hazard functions to be accommodated, and these have been seen to generate Phillips curves in which lagged inflation has a positive coefficient.

The estimation results for the parameters α, φ_1 and β of the first-order recursive model are displayed in Table 2. What is immediately apparent is that the estimates of φ_1 are positive and statistically significant at the 5% level for all specifications. This represents a strong rejection of the Calvo model, which is equivalent to the null hypothesis $\varphi_1 = 0$ within this class of models. By invoking the results of Proposition 1, the point estimates of φ_1 imply a hazard function that is increasing everywhere. The parameter α now needs to be interpreted differently from the Calvo model results in Table 1. Here it is merely the probability of a firm changing a price that was posted at some time during the previous quarter. The estimates of α in Table 2 are much lower than those in Table 1, and are not significantly different from zero in any case. In the other columns of Table 2, the estimates of β are quite low but insignificantly different from one. The J-statistics fail to reject the over-identifying restrictions in any specification.

As discussed in section 4.1, the restrictions on the parameters α and φ_1 needed to ensure that the implied hazard function is well defined are not imposed at the estimation stage. The rationale for doing this is to allow these theoretical restrictions to be tested and thus assess whether inflation dynamics are consistent with a well-defined hazard function model. Proposition 3 shows that for the first-order model, the restrictions are given by the inequalities in (9). The first of these is $0 < \alpha < 1$. The estimate from the preferred specification passes this test, and while the point estimates under the alternative specifications fail to do so, none of the violations is statistically significant. The next condition from (9) to check is given as $(1-\alpha)^2/4-\varphi_1$ in a column of Table 2. According to (9), this should be positive if the hazard function is to be well defined everywhere. There are small and statistically insignificant violations of this condition in three out of the four specifications considered in Table 2. The third condition required by (9) is automatically satisfied because all the point estimates of φ_1 are positive.

Plots of the implied hazard and survival functions for the estimated first-order recursive model are shown in Figure 3. As usual, these are plotted for the parameters estimated under the preferred specification, that is, the first row of Table 2. The hazard function is upward sloping because the point estimate of φ_1 is positive. The one- and two-standard-deviation bands in Figure 3 imply that the estimated hazard function starts at a point insignificantly different from zero (given by the parameter α) for prices that have been changed very recently, and rises to a point insignificantly different from one for prices that have been left fixed for six or seven quarters. Because the estimated parameters fail to satisfy all the inequalities in (9), the point estimate of the hazard function is not well defined beyond seven quarters. The bands also become very wide as the duration of price stickiness increases.[15] In spite of the wide bands,

[15]This should not be surprising. Once the proportion of firms using a price of a particular age shrinks to zero

hypotheses about the slope of the hazard function can be tested directly using the sign of the parameter φ_1 as discussed above.

The coefficients of the implied Phillips curve for the first-order model are given in Table 3. The key point to note is that the significantly positive φ_1 parameter from Table 2 translates into a significantly positive coefficient on inflation lagged one quarter. Thus the estimated hazard function shown in Figure 3 implies a non-negligible amount of intrinsic inflation persistence, comparable in magnitude to that found by Galí and Gertler (1999) for a model with backward-looking firms. The parameter estimates also imply a significantly negative coefficient on expected inflation two quarters in the future, but the coefficient on inflation one quarter ahead remains significantly positive. As has been discussed, this negative coefficient should be interpreted merely as a reduction of the weight attached to future inflation in determining current inflation. In the preferred specification, the coefficient on unit labour costs is positive and statistically significant at the 5% level.

It is interesting to note that the estimated first-order model is able to generate intrinsic inflation persistence without requiring noticeably more price stickiness than is found in the estimated Calvo model. While the expected probability of price adjustment α^e is estimated to be larger in the Calvo model, Tables 1 and 2 show that the average duration $\mathcal{D}^e$ of price stickiness is 2.198 quarters for the first-order model and 2.469 quarters for the Calvo model. Inspection of the hazard functions in Figures 2 and 3 shows that the hazard function for the first-order model is above that of the Calvo model for all prices except those posted in the previous quarter.

On the basis of the results for the first-order recursive model, a flat hazard function is clearly rejected by the data in favour of an alternative with a monotonically increasing hazard function. This new model offers a more promising account of inflation dynamics. However, restricting attention to hazard functions generated by a first-order recursion in (7) still imposes essentially arbitrary limitations on the range of allowed hazard function shapes. For this reason, it is desirable to consider higher-order models. Proposition 2 shows that if the true hazard function satisfies Assumption 1 and if n is made sufficiently large then a recursion of the form (7) is able to approximate the model as accurately as is required. But econometric practicalities put some limits on the maximum order of model that can be estimated because the number of terms in the Phillips curve (23) rises in step with the order of the recursion.

Starting from $n = 2$, progressively higher orders of recursion (7) were estimated. The extra parameter φ_2 introduced by the second-order model turns out to be statistically insignificant. More success is had with the cases $n = 3$ and $n = 4$ where both φ_3 and φ_4 are highly significant. Beyond that, no additional statistically significant slope parameters are found, with models up to $n = 8$ being estimated. The full set of results is not reported here owing to limited space, but the fourth-order model is presented as typical of these findings. The estimates of the parameters

the probability of such a price being changed ceases to be identified. The Calvo model avoids this problem by asserting the probability is the same for prices of all ages.

α, $\{\varphi_i\}_{i=1}^4$ and β are displayed in Table 4.

The initial probability of price adjustment α is estimated to be low and insignificantly different from zero in all specifications. The slope parameters φ_1 and φ_4 are found to be significantly positive; φ_2 is insignificantly different from zero, and φ_3 is significant and negative. The mixture of positives and negatives suggests that the implied hazard function is no longer monotonic. This is confirmed by the plots of the hazard and survival functions in Figure 4. The point estimate of the hazard function begins at a point insignificantly different from zero for prices that have just been set, and begins to rise during the first and second quarters of a spell of price stickiness. It falls back in the third quarter and then rises again in the fourth quarter. At the beginning of the second year of price stickiness the hazard function rises sharply, but again falls back somewhat after the middle of the second year. Finally, it rises very sharply again after around two years of price stickiness, reaching a level not significantly different from one. These results suggest that firms are more likely to make a price adjustment around the first and second anniversaries of their previous price change: a feature that also finds some support in the micro evidence. After a duration of two years, the hazard function is estimated too imprecisely to draw any firm conclusions.

As is the case with the estimated first-order model, the point estimate of the fourth-order hazard function in Figure 4 is well defined for most, but not all, durations of price stickiness. The most prominent failure is in the fourth quarter where the point estimate dips below zero, but this deviation is not statistically significant. The estimates of the hazard function after the eighth quarter are much too loose to be able to detect any statistically significant deviation here either. Therefore, the estimated fourth-order model is very close to implying a well-defined hazard function, and no statistically significant failure to meet this requirement can be found.

The implied Phillips curve for the estimated fourth-order model is exhibited in Table 5. The first and the fourth lags of inflation now have significantly positive coefficients, the second lag's coefficient is a small and insignificant positive number, and the third lag has a negative coefficient that is statistically significant. Overall, the positive coefficients on the lags of inflation clearly dominate. Thus the hazard function in Figure 4 provides a rationale for why inflation rates one quarter ago and one year ago contribute positively to intrinsic inflation persistence. Among the coefficients on future inflation there is a mixture of positives and negatives. In the preferred specification, the coefficient of unit labour costs remains positive and significant at the 10% level.

The conclusions drawn from these results are: that a flat hazard function is resoundingly rejected; that hazard functions with upward sloping sections are found; that these can justify the quantitative importance of the positive coefficients of lagged inflation found for estimated Phillips curves; that although the estimated hazard functions are not well defined for all durations of price stickiness, no statistically significant rejection of a well-defined hazard function is found.

4.3 Comparison with the microeconometric evidence

There are now many studies that estimate the hazard function for price adjustment using microeconomic data on individual prices (Baumgartner et al., 2005; Campbell and Eden, 2005; Cecchetti, 1986; Dias et al., 2005; Fougère et al., 2005; Götte et al., 2005; Nakamura and Steinsson, 2007). There are a range of findings in this literature. Papers such as Götte et al. (2005) and Cecchetti (1986) find strong evidence in favour of an upward-sloping hazard function. Both papers use a small number of goods, but have data spanning several decades. Others such as Campbell and Eden (2005) and Dias et al. (2005) find strong evidence that the hazard function is downward sloping. Some studies such as Baumgartner et al. (2005) agree that the hazard function is generally downward sloping, but find that the negative slope is interrupted by sharp spikes at regular intervals. This group of studies uses data on a very large number of goods, but these data are drawn from a relatively small number of years in the last decade.

It is argued that some of the findings of downward-sloping hazards can be explained as the result of a heterogeneity bias (Álvarez et al., 2005). Many studies include a wide range of products that have different degrees of price stickiness. Heterogeneity biases estimates of the hazard function slope downward because goods with more flexible prices are less likely to be found to have long spells of price stickiness.[16] As a result of this criticism, studies such as Nakamura and Steinsson (2007) and Fougère et al. (2005) take careful steps to control for heterogeneity. Nakamura and Steinsson (2007) allow the level of each product's hazard function to be different. They find that the estimated hazard function is then largely flat, with a large spike after one year. Fougère et al. (2005) allow both the level and the slope of the hazard function to differ across products. The results are now mixed, with a range of increasing and decreasing hazards found for different goods and different types of retail outlet. But increasing hazard functions are in the majority.

If the results based on macro data from section 4 are compared only with those microeconometric studies that span several decades including times of high as well as low inflation, then there is no contradiction between the micro and macro evidence. However, the micro studies in this group draw on a rather narrow range of goods, though on the other hand this narrow range may also be a virtue if the heterogeneity bias is thought to be a serious problem. Evidence from the more comprehensive micro studies working with data from the 1990s and 2000s does present a prima facie contradiction to the macro-data estimates of hazard functions for the period 1960–2003 derived in this paper. There are two points to bear in mind here. First, it remains to be seen how robust the finding of a downward-sloping hazard function is once heterogeneity is properly controlled for. Second, the theoretical case for an upward-sloping hazard function is strongest in periods of higher inflation, hence the hazard function slope may not be a structural feature of the economy. Thus the failure to find a positive slope using data only

[16] See Heckman and Singer (1984) and Kiefer (1988) for more discussion of the problem of heterogeneity in duration analysis.

from a period of low and stable inflation may not be surprising.[17] Finally, it should be noted that hazard function "spikes" found in the microeconometric literature are of course evidence for a sharply upward-sloping section of the hazard function, and thus contribute to explaining intrinsic inflation persistence. The fact that the spikes also imply a sharply downward-sloping section does not offset this effect because the existence of a spike at some duration implies that much fewer price spells survive beyond that duration to where the hazard function is actually downward sloping.

5. Conclusions

This paper has studied the link between intrinsic inflation persistence and the price-setting behaviour of firms. Intrinsic inflation persistence refers to inflation that occurs purely as a result of past pricing decisions and cannot be explained by current and expected future fundamentals such as unit labour costs, output gaps, monetary policy, or cost-push shocks. When intrinsic inflation persistence is present in an economy, it is not possible for the central bank to bring inflation immediately back to target without some loss of output, even if the shocks that gave rise to the inflation have dissipated. Most empirical studies conclude that inflation determination is not a purely forward-looking process and that inflation contains a significant backward-looking component, though the reasons for the existence of this intrinsic inflation persistence are considered to be a puzzle. But the results of this paper show that there is no contradiction between such persistence and profit-maximizing, forward-looking price setting by firms.

What turns out to be important for intrinsic inflation persistence is not how much price stickiness there is on average, but whether there are systematic differences between the stickiness of prices of different ages. In particular, newer prices need to be stickier than older prices in order to explain this persistence. If older prices are more likely to be adjusted then the "catch-up" effect of firms whose prices have remained fixed for a long time has a larger impact on current inflation than the "roll-back" effect of firms who have recently adjusted their prices after a shock, leading to sustained rises in the price level even following temporary shocks.

In terms of the hazard function for price changes, the most important feature influencing intrinsic inflation persistence is the slope, not the level. The level could be very high or very low, representing the extremes of price flexibility or stickiness, but as long as the hazard function remains flat there can be no intrinsic persistence. To provide a rationale for the type of intrinsic inflation persistence described above, where high inflation in the past makes it harder to achieve low inflation today without sacrificing output, it is necessary that the hazard function is predominantly upward sloping. If the hazard function were predominantly downward sloping then a perverse result is obtained whereby the higher inflation has been in the past, the

[17] A preliminary subsample analysis (not reported here) suggests that macro-based estimates of the hazard function using only data from the mid 1980s to 2000s find a hazard function that is initially downward sloping and followed by less steeply upward-sloping sections.

easier it is to achieve low inflation today without cost.

In order to study the link between the hazard function for price changes and intrinsic inflation persistence, a methodological innovation is introduced that shows how essentially all hazard functions have a recursive representation. This allows much simpler expressions to be obtained for the Phillips curves implied by non-constant hazard models of price setting. The resulting class of simple Phillips curves closely resembles the "hybrid" New Keynesian Phillips curves widely used in empirical work and policy analysis, but which are thought to have weak theoretical foundations because they introduce lags of inflation arbitrarily that are not present in the standard New Keynesian model. But there is nothing arbitrary about these lags: they are simply the implication of profit-maximizing behaviour by firms when the likelihood of adjusting prices depends on the amount of time elapsed since the previous price change. By using this new class of Phillips curves, intrinsic inflation persistence can be precisely defined in terms of the coefficients of lagged inflation. It is shown analytically how positively sloped hazard functions imply positive values of these coefficients and negatively sloped hazards imply that these coefficients have negative values.

By building on these insights, this paper then sets out a method for estimating the hazard function for price changes without needing microeconomic data on individual prices. The hazard function can be identified and estimated using only macroeconomic time-series on inflation and unit labour costs and simple econometric techniques. The results of this exercise strongly reject the flat hazard function underlying the standard New Keynesian Phillips curve. The point estimates suggest that the hazard function is predominantly upward sloping, starting from a probability of price adjustment not significantly different from zero for prices that have been changed very recently, rising to a probability not significantly different from one for prices that have remained sticky for eight or more quarters. In between, the hazard function initially rises and then falls back during the first year of price stickiness. There is then a sudden increase in the probability of price adjustment towards the end of the first year and the beginning of the second year. A similar sudden rise is also found towards the end of the second year. No statistically significant rejection of a well-defined hazard function is found. Therefore, in summary, an upward-sloping hazard function model offers a significantly better account of empirical inflation dynamics without introducing any arbitrary backward-looking behaviour.

Table 1

Estimates of the Calvo pricing model ($n = 0$)

Specification[§]	α	β	$\alpha^{e\sharp}$	$\mathcal{D}^{e\sharp}$	J-stat[‡]	$\mathbb{E}_t\pi_{t+1}{}^{\natural}$	$\mathsf{x}_t{}^{\natural}$
$\eta_{cx} = 0.4$	0.405^{**}	0.966^{**}	0.405^{**}	2.469^{**}	17.223	0.966^{**}	0.116^{**}
$\mathcal{N}(1)$	(0.053)	(0.028)	(0.053)	(0.322)	[0.575]	(0.028)	(0.040)
$\eta_{cx} = 0.4$	0.217^{**}	0.985^{**}	0.217^{**}	4.601^{**}	14.804	0.985^{**}	0.025
$\mathcal{N}(2)$	(0.106)	(0.024)	(0.106)	(2.246)	[0.735]	(0.024)	(0.028)
$\eta_{cx} = 1.0$	0.226^{**}	0.975^{**}	0.226^{**}	4.432^{**}	15.873	0.975^{**}	0.071^{**}
$\mathcal{N}(1)$	(0.046)	(0.026)	(0.046)	(0.895)	[0.666]	(0.026)	(0.033)
$\eta_{cx} = 1.0$	0.142^{*}	0.985^{**}	0.142^{*}	7.064^{*}	14.804	0.985^{**}	0.025
$\mathcal{N}(2)$	(0.073)	(0.024)	(0.073)	(3.637)	[0.735]	(0.024)	(0.028)

Notes: Estimation of the parameters α and β is by GMM using U.S. quarterly data 1960:Q1–2003:Q4. The moment conditions are given in (29) and the data are described in section 4.1. 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.

[§] Each specification is a calibrated value of the parameter η_{cx} and a choice of either normalization $\mathcal{N}(1)$ or $\mathcal{N}(2)$. See section 4.1 for further details.

[♯] The expected probability of price adjustment α^e and the expected duration of price stickiness $\mathcal{D}^e$ are inferred from the estimated parameters using equation (16).

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

[♮] These are the implied coefficients of the New Keynesian Phillips curve in (25).

Table 2
Estimates of the first-order recursive model ($n = 1$)

Specification[§]	α	φ_1	β	$\alpha^{e\sharp}$	$\mathcal{D}^{e\sharp}$	$\frac{1}{4}(1-\alpha)^2 - \varphi_1$[¶]	J-stat[‡]
$\eta_{cx} = 0.4$	0.132	0.222^{**}	0.899^{**}	0.354^{**}	2.198^{**}	−0.034	11.589
$\mathcal{N}(1)$	(0.104)	(0.063)	(0.075)	(0.057)	(0.303)	(0.031)	[0.868]
$\eta_{cx} = 0.4$	−0.205	0.406^{**}	0.913^{**}	0.201	2.955^{**}	−0.043	9.578
$\mathcal{N}(2)$	(0.267)	(0.155)	(0.130)	(0.128)	(1.361)	(0.048)	[0.945]
$\eta_{cx} = 1.0$	−0.048	0.265^{**}	0.926^{**}	0.217^{**}	3.389^{**}	0.010	10.862
$\mathcal{N}(1)$	(0.107)	(0.081)	(0.067)	(0.047)	(0.693)	(0.038)	[0.900]
$\eta_{cx} = 1.0$	−0.340	0.471^{**}	0.913^{**}	0.131	4.032^{*}	−0.022	9.578
$\mathcal{N}(2)$	(0.231)	(0.163)	(0.130)	(0.088)	(2.129)	(0.048)	[0.945]

Notes: Estimation of the parameters α, φ_1 and β is by GMM using U.S. quarterly data 1960:Q1–2003:Q4. See the notes to Table 1 for further details.

¶ This value should be positive to ensure that the upward-sloping hazard functions generated by a first-order recursion are well defined. See Proposition 3 and equation (9) for more details.

Table 3
Implied Phillips curve for the estimated first-order recursive model ($n = 1$)

Specification[§]	π_{t-1}	$\mathbb{E}_t\pi_{t+1}$	$\mathbb{E}_t\pi_{t+2}$	x_t
$\eta_{cx} = 0.4$	0.271^{**}	0.923^{**}	-0.219^{**}	0.069^{**}
$\mathcal{N}(1)$	(0.051)	(0.056)	(0.036)	(0.031)
$\eta_{cx} = 0.4$	0.326^{**}	0.939^{**}	-0.272^{**}	0.015
$\mathcal{N}(2)$	(0.049)	(0.093)	(0.061)	(0.024)
$\eta_{cx} = 1.0$	0.255^{**}	0.944^{**}	-0.219^{**}	0.053^{*}
$\mathcal{N}(1)$	(0.053)	(0.051)	(0.041)	(0.028)
$\eta_{cx} = 1.0$	0.326^{**}	0.939^{**}	-0.272^{**}	0.015
$\mathcal{N}(2)$	(0.049)	(0.093)	(0.061)	(0.024)

Notes: This table reports the coefficients of the Phillips curve (27) implied by the parameter estimates in Table 2 for each specification. Standard errors are given in parentheses and are calculated using the delta method. See the notes to Tables 1 and 2 for more details about the estimation method.

Table 4

Parameter estimates of the fourth-order recursive model (n = 4)

Specification§	α	φ_1	φ_2	φ_3	φ_4	β	$\alpha^{e\sharp}$	$\mathcal{D}^{e\sharp}$	J-stat‡
$\eta_{cx} = 0.4$	0.148	0.165**	−0.013	−0.202**	0.187**	0.866**	0.285**	2.528**	6.110
$\mathcal{N}(1)$	(0.110)	(0.073)	(0.066)	(0.079)	(0.048)	(0.072)	(0.069)	(0.432)	[0.978]
$\eta_{cx} = 0.4$	0.072	0.205**	−0.029	−0.171**	0.177**	1.045**	0.254**	2.577**	4.671
$\mathcal{N}(2)$	(0.163)	(0.086)	(0.067)	(0.077)	(0.058)	(0.067)	(0.093)	(0.599)	[0.995]
$\eta_{cx} = 1.0$	0.001	0.187**	−0.029	−0.208**	0.221**	0.876**	0.173**	3.536**	5.896
$\mathcal{N}(1)$	(0.107)	(0.084)	(0.071)	(0.089)	(0.055)	(0.074)	(0.051)	(0.725)	[0.981]
$\eta_{cx} = 1.0$	−0.058	0.233**	−0.036	−0.171**	0.207**	1.045**	0.175**	3.000**	4.671
$\mathcal{N}(2)$	(0.159)	(0.097)	(0.072)	(0.084)	(0.068)	(0.067)	(0.068)	(0.932)	[0.995]

Notes: Estimation of the parameters α, $\{\varphi_i\}_{i=1}^4$ and β is by GMM using U.S. quarterly data 1960:Q1–2003:Q4. See the notes to Table 1 for further details.

Table 5

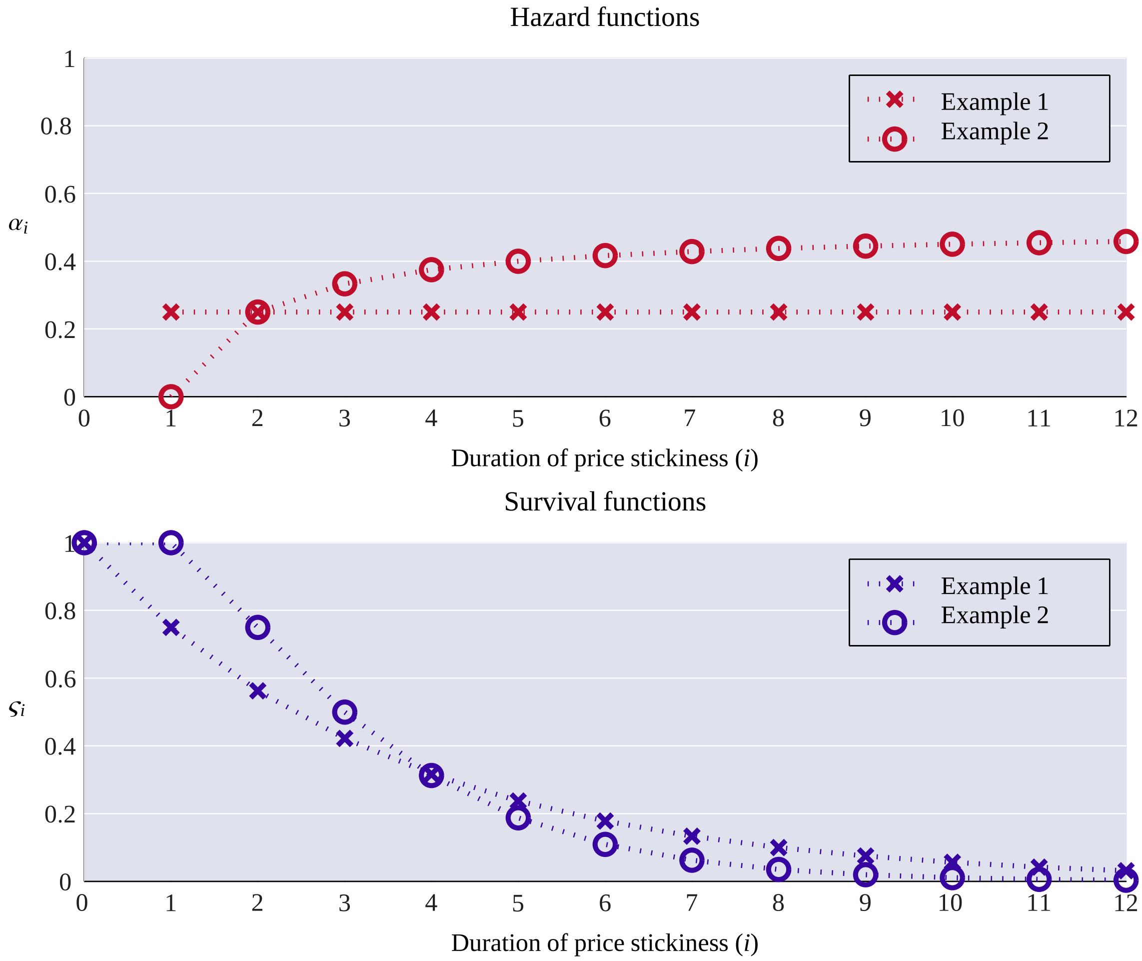
Implied Phillips curve for the estimated fourth-order recursive model (n = 4)

Specification[§]	π_{t-1}	π_{t-2}	π_{t-3}	π_{t-4}	$\mathbb{E}_t\pi_{t+1}$	$\mathbb{E}_t\pi_{t+2}$	$\mathbb{E}_t\pi_{t+3}$	$\mathbb{E}_t\pi_{t+4}$	$\mathbb{E}_t\pi_{t+5}$	x_t
η_{cx} = 0.4	0.183**	0.008	−0.182**	0.222**	0.889**	−0.136**	−0.004	0.119**	−0.108**	0.048*
$\mathcal{N}(1)$	(0.067)	(0.065)	(0.080)	(0.058)	(0.061)	(0.058)	(0.046)	(0.058)	(0.047)	(0.028)
η_{cx} = 0.4	0.207**	0.012	−0.172**	0.184**	1.034**	−0.228**	−0.014	0.195**	−0.229**	0.025
$\mathcal{N}(2)$	(0.064)	(0.070)	(0.074)	(0.057)	(0.050)	(0.071)	(0.077)	(0.091)	(0.086)	(0.023)
η_{cx} = 1.0	0.181**	0.005	−0.179**	0.220**	0.898**	−0.138**	−0.002	0.122**	−0.114**	0.041
$\mathcal{N}(1)$	(0.067)	(0.065)	(0.080)	(0.058)	(0.063)	(0.059)	(0.048)	(0.060)	(0.050)	(0.027)
η_{cx} = 1.0	0.207**	0.012	−0.172**	0.184**	1.034**	−0.228**	−0.014	0.195**	−0.229**	0.025
$\mathcal{N}(2)$	(0.064)	(0.070)	(0.074)	(0.057)	(0.050)	(0.071)	(0.077)	(0.091)	(0.086)	(0.023)

Notes: This table reports the coefficients of the Phillips curve (23) implied by the parameter estimates in Table 4 for each specification. Standard errors are given in parentheses and are calculated using the delta method. See the notes to Tables 1 and 4 for more details about the estimation.

Figure 1

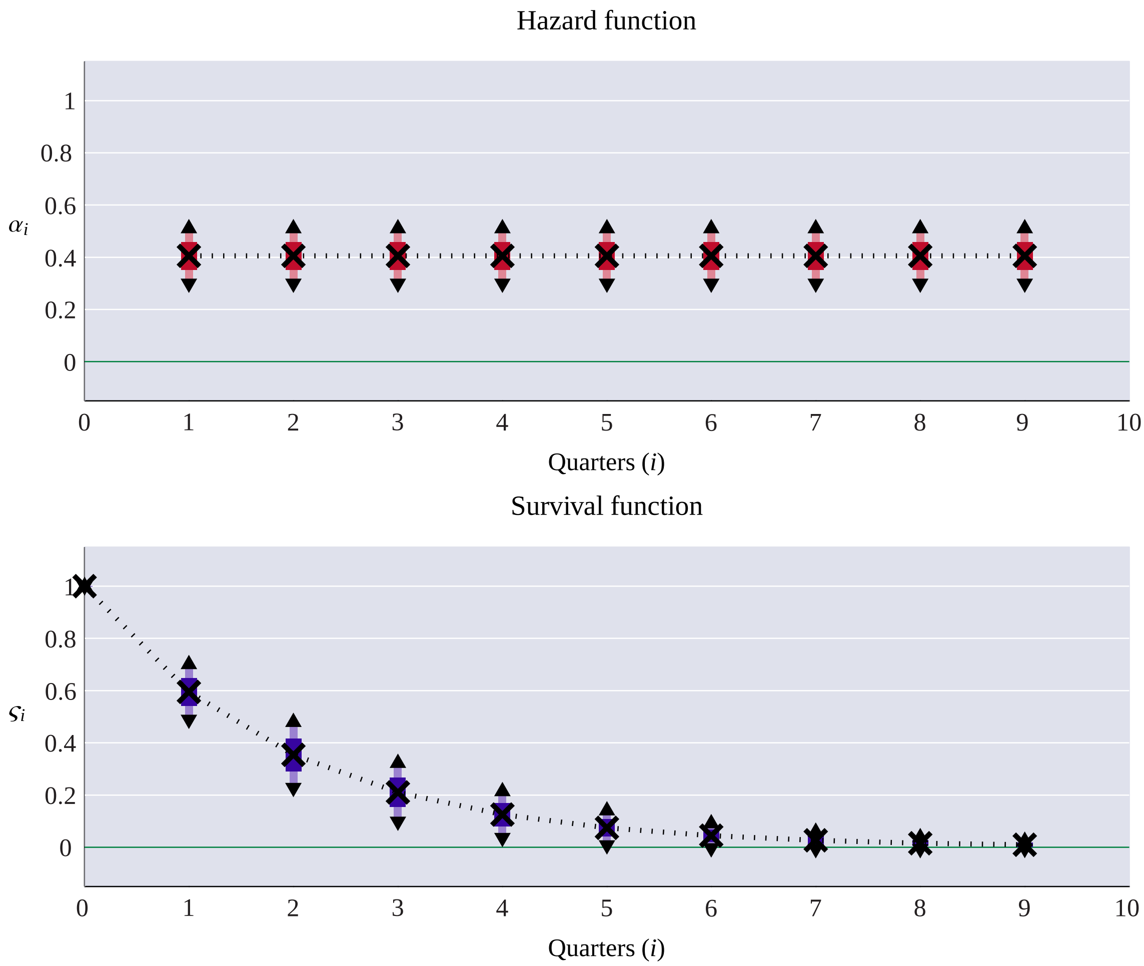
Comparison of the hazard and survival functions implied by Examples 1 and 2



Notes: The hazard function sequences $\{\alpha_i\}_{i=0}^{\infty}$ are generated by equation (7). Example 1 sets $n = 0$ and $\alpha = 0.25$; Example 2 uses $n = 1$, $\alpha = 0$ and $\varphi = 0.25$. The corresponding survival function sequences $\{\varsigma_i\}_{i=0}^{\infty}$ are obtained from (8).

Figure 2

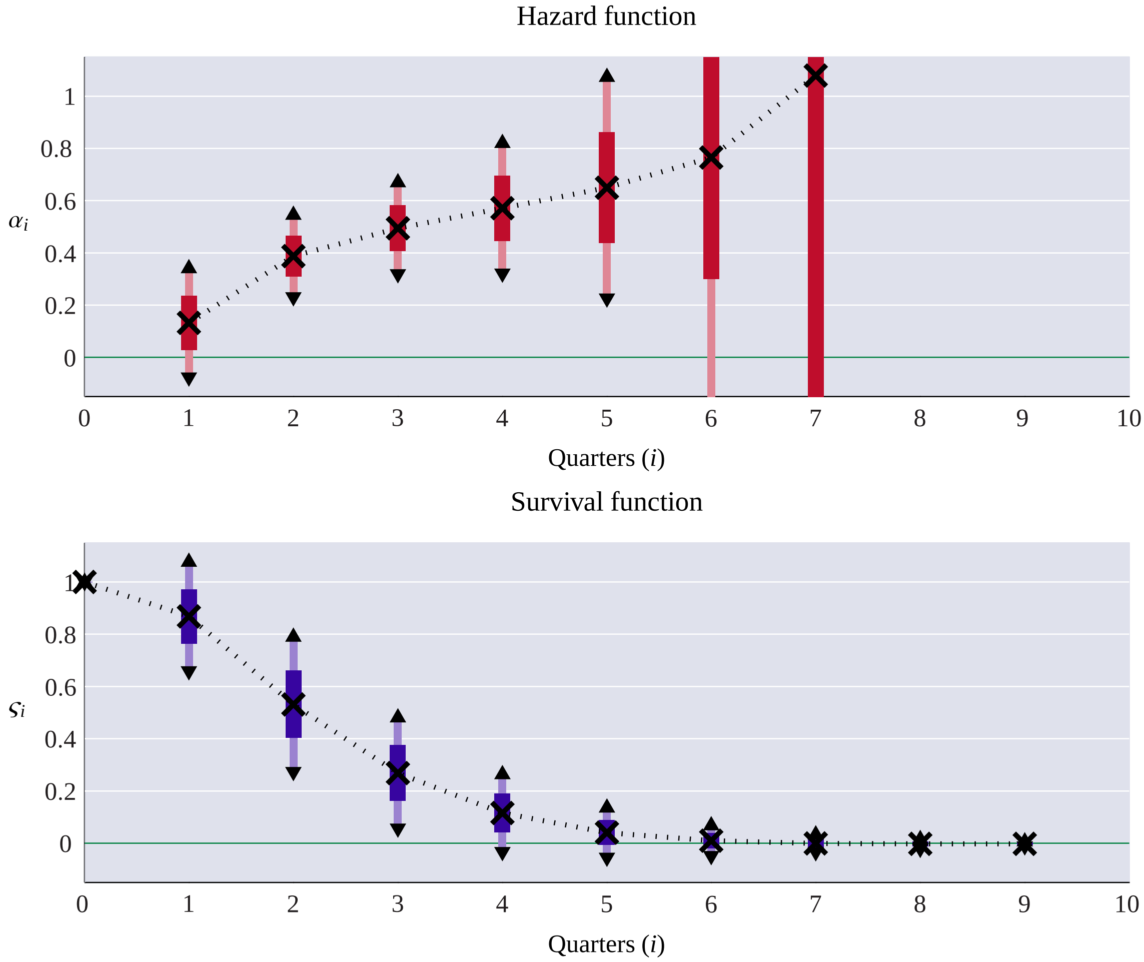
Estimated hazard function and survival function for the Calvo pricing model ($n = 0$)



Notes: The hazard and survival functions are calculated from (7) and (8) using the estimated parameters from Table 1 under the preferred specification $\eta_{cx} = 0.4$ and $\mathcal{N}(1)$. In the graphs, the point estimate is shown as a cross in the middle of a darkly shaded thick bar that represents a one-standard-deviation band around the point estimate, which is itself surrounded by a lightly shaded thin bar giving the two-standard-devations band. The standard deviation is calculated using the delta method.

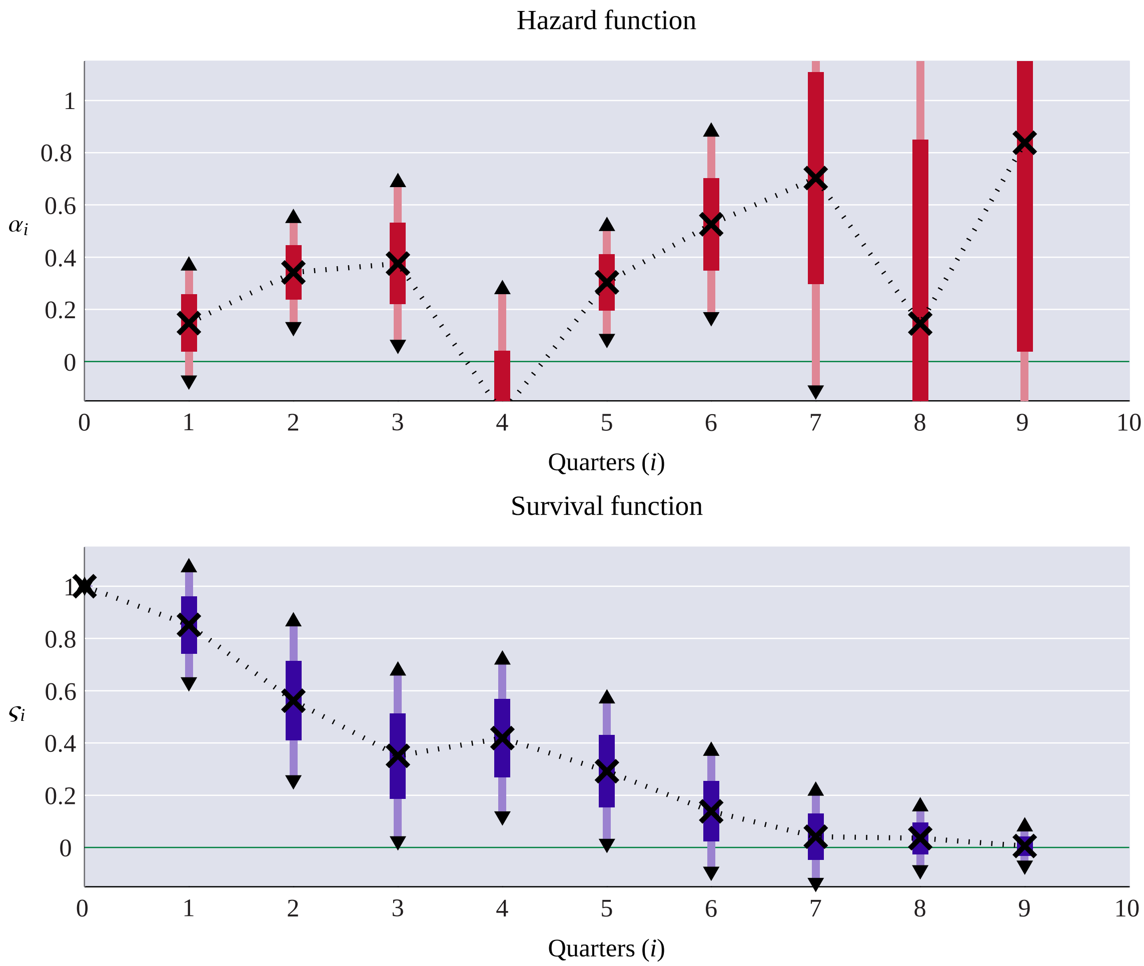
Figure 3

Estimated hazard function and survival function for the first-order recursive model (n = 1)



Notes: The hazard and survival functions are calculated from (7) and (8) using the estimated parameters from Table 1 under the preferred specification η_{cx} = 0.4 and $\mathcal{N}(1)$. In the graphs, the point estimate is shown as a cross in the middle of a darkly shaded thick bar that represents a one-standard-deviation band around the point estimate, which is itself surrounded by a lightly shaded thin bar giving the two-standard-devations band. The standard deviation is calculated using the delta method.

Figure 4
Estimated hazard function and survival function for the fourth-order recursive model ($n = 4$)



Notes: The hazard and survival functions are calculated from (7) and (8) using the estimated parameters from Table 1 under the preferred specification $\eta_{cx} = 0.4$ and $\mathcal{N}(1)$. In the graphs, the point estimate is shown as a cross in the middle of a darkly shaded thick bar that represents a one-standard-deviation band around the point estimate, which is itself surrounded by a lightly shaded thin bar giving the two-standard-devations band. The standard deviation is calculated using the delta method.

A. Technical appendix

A.1 Proof of Lemma 1

Let $\mathfrak{g}$ be a number lying strictly between $(1-\mathfrak{d})$ and 1, which must satisfy $0 < \mathfrak{g} < 1$. The definition of $\mathfrak{g}$ then implies that $\varsigma_{i+1} \leq \mathfrak{g}\varsigma_i$ and hence $\varsigma_i \leq \mathfrak{g}^i \varsigma_0$ for all $i \geq 0$. Now let $\mathfrak{x}$ be any number strictly between 1 and $\mathfrak{g}^{-1}$.

(i) The constant $\mathfrak{x}$ is defined so that $1 < \mathfrak{x} < \mathfrak{g}^{-1}$ which implies $0 < \mathfrak{g}\mathfrak{x} < 1$ and hence that $\sum_{i=0}^{\infty} \mathfrak{g}^i |z|^i$ converges for all z in the disc $\mathfrak{D}_\mathfrak{x}$. By applying the triangle inequality to the function $\varsigma(z)$ it follows that $|\varsigma(z)| \leq \varsigma_0/(1-\mathfrak{g}\mathfrak{x})$ if $z \in \mathfrak{D}_\mathfrak{x}$, and therefore that $\varsigma(z)$ is analytic on $\mathfrak{D}_\mathfrak{x}$.

(ii) Construct a new function $\mathcal{G}(z) \equiv (1-\mathfrak{g}z)\varsigma(z)$, which inherits the property that it is analytic on $\mathfrak{D}_\mathfrak{x}$ from $\varsigma(z)$. Using the definition of $\varsigma(z)$, $\mathcal{G}(z)$ can be written as the sum of $\mathcal{G}_0(z) \equiv \varsigma_0$ and $\mathcal{G}_1(z) \equiv -\sum_{i=1}^{\infty}(\mathfrak{g}\varsigma_{i-1} - \varsigma_i)z^i$. Now take any $z \in \mathfrak{D}_\mathfrak{x}$. Since $|z|^i \leq \mathfrak{x}^i$ and $\varsigma_i \leq \mathfrak{g}\varsigma_{i-1}$ the triangle inequality and some algebraic manipulations show that $|\mathcal{G}_1(z)| \leq \mathfrak{g}\mathfrak{x}\varsigma_0 - \sum_{i=1}^{\infty}(1-\mathfrak{g}\mathfrak{x})\mathfrak{x}^i\varsigma_i$. Because $0 < \mathfrak{g}\mathfrak{x} < 1$ and $|\mathcal{G}_0(z)| = \varsigma_0$, it is established that $|\mathcal{G}_1(z)| < |\mathcal{G}_0(z)|$ for all $z \in \mathfrak{D}_\mathfrak{x}$.

As a constant function, $\mathcal{G}_0(z)$ must be analytic, and $\mathcal{G}_1(z)$ inherits the property from $\mathcal{G}(z)$. Since $\mathcal{G}(z) = \mathcal{G}_0(z) + \mathcal{G}_1(z)$, Rouché's Theorem implies that $\mathcal{G}(z)$ and $\mathcal{G}_0(z)$ have the same number of zeros on $\mathfrak{D}_\mathfrak{x}$.[18] Since $\mathcal{G}_0(z)$ clearly has no zeros on this set, neither has $\mathcal{G}(z)$. Because its definition ensures that $\mathcal{G}(z)$ inherits any roots of $\varsigma(z) = 0$, this precludes $\varsigma(z)$ having a zero in $\mathfrak{D}_\mathfrak{x}$ as well.

A.2 Proof of Proposition 1

When deriving the following results, it is convenient to assume $n = \infty$ in (7) and set any superfluous φ_i parameters to zero.

(a) This follows immediately from equation (7) by induction.

(b) i. Suppose that the first h recursive parameters are non-negative. Equation (7) then clearly implies $\alpha_2 \geq \alpha_1$. Now suppose that $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{i-1} \leq \alpha_i$ has already been established for some $i \leq h$. The slope of the hazard function can be obtained from (7):

$$\alpha_{i+1} - \alpha_i = \sum_{j=1}^{i-1} \varphi_j(\alpha_i - \alpha_{i-j}) \left(\prod_{k=i-j}^{i} (1-\alpha_k) \right)^{-1} + \varphi_i \left(\prod_{k=1}^{i} (1-\alpha_k) \right)^{-1} \tag{A.1}$$

Given the initial supposition, the second term is unambiguously positive, and assuming the truth of the induction step, the first term must be non-negative. Hence $\alpha_{i+1} \geq \alpha_i$, and so the claim is proved by induction.

ii. Now suppose that $\alpha_{i+1} > \alpha_i$, but $\varphi_i \leq 0$ for all $i = 1, \ldots, h$. By making use of (A.1) this leads to a contradiction, thus proving the claim.

iii. This follows from the result in part (c)i. below.

[18] I am indebted to Graham Brightwell for suggesting this use of Rouché's Theorem. See any text on complex analysis, such as Gamelin (2001), for further details about the theorem.

(c) i. Define the terms of the sequence $\{\phi_i\}_{i=1}^{\infty}$ to be $\phi_1 \equiv 1-\alpha$ and $\phi_i \equiv -\varphi_{i-1}$ for $i \geq 2$. With these definitions, the survival function recursion (8) reduces to $\varsigma_i = \sum_{j=1}^{i} \phi_j \varsigma_{i-j}$, which can be flipped around to yield $\phi_i = \varsigma_i - \sum_{j=1}^{i-1} \phi_j \varsigma_{i-j}$ since $\varsigma_0 = 1$. Now define another sequence $\{\mathfrak{a}_i\}_{i=1}^{\infty}$ using $\mathfrak{a}_1 \equiv 1-\alpha$ and $\mathfrak{a}_i \equiv \alpha_{i-1} - \alpha_i$ for $i \geq 2$. With these definitions and equation (6) it follows that $\varsigma_i = \prod_{j=1}^{i} \left(\sum_{k=1}^{j} \mathfrak{a}_k\right)$. Next, define some sets $\mathcal{J}_i \equiv \left\{ (\ell_1, \ldots, \ell_i) \,\middle|\, \ell_j \in \mathbb{N} \;,\; 1 \leq \ell_j \leq j \right\}$, and use these to reverse the order of summation and multiplication in the expression for ϕ_i, resulting in $\varsigma_i = \sum_{(\ell_1,\ldots,\ell_i)\in\mathcal{J}_i} \prod_{h=1}^{i} \mathfrak{a}_{\ell_h}$

Now construct new sets $\mathcal{J}_i^* \equiv \mathcal{J}_i \backslash \left(\bigcup_{j=1}^{i-1} \left(\mathcal{J}_j^* \times \mathcal{J}_{i-j}\right)\right)$ recursively from the sets $\mathcal{J}_i$. The recursion is initialized with $\mathcal{J}_1^* \equiv \mathcal{J}_1 \equiv \{\ (1)\ \}$. It is straightforward to check that these definitions lead to well-defined sets since $\mathcal{J}_j^*$ is always a subset of $\mathcal{J}_j$.

It is claimed that $\phi_i = \sum_{(\ell_1,\ldots,\ell_i)\in\mathcal{J}_i^*} \prod_{h=1}^{i} \mathfrak{a}_{\ell_h}$. By looking at the definitions of $\mathcal{J}_1^*$, $\mathfrak{a}_1$ and ϕ_1, this claim is certainly true for $i = 1$. Suppose it has already been proved for $i = 1, \ldots, j-1$. Given the induction step and the earlier expression for ς_j:

$$\phi_j = \left(\sum_{(\ell_1,\ldots,\ell_j)\in\mathcal{J}_j} \prod_{h=1}^{j} \mathfrak{a}_{\ell_h}\right) - \sum_{k=1}^{j-1} \left(\sum_{(\ell_1,\ldots,\ell_j)\in(\mathcal{J}_k^*\times\mathcal{J}_{j-k})} \prod_{h=1}^{j} \mathfrak{a}_{\ell_h}\right) \tag{A.2}$$

It follows that $\phi_j = \sum_{(\ell_1,\ldots,\ell_j)\in\mathcal{J}_j\backslash\left(\bigcup_{k=1}^{j-1}(\mathcal{J}_k^*\times\mathcal{J}_{j-k})\right)} \left(\prod_{h=1}^{j} \mathfrak{a}_{\ell_h}\right)$ if the sets $\mathcal{J}_k^* \times \mathcal{J}_{j-k}$ and $\mathcal{J}_l^* \times \mathcal{J}_{j-l}$ are disjoint for all $k \neq l$, which would prove the earlier claim.

Now suppose for contradiction that $(\mathcal{J}_k^* \times \mathcal{J}_{j-k}) \cap (\mathcal{J}_l^* \times \mathcal{J}_{j-l}) \neq \emptyset$, and without loss of generality take $k > l$. Hence there is a vector $(\ell_1, \ldots, \ell_j) \in \mathbb{N}^j$ such that $(\ell_1, \ldots, \ell_k) \in \mathcal{J}_k^*$, $(\ell_1, \ldots, \ell_l) \in \mathcal{J}_l^*$, and $(\ell_{l+1}, \ldots, \ell_j) \in \mathcal{J}_{j-l}$. This implies that $(\ell_{l+1}, \ldots, \ell_k) \in \mathcal{J}_{k-l}$ and hence $(\ell_1, \ldots, \ell_k) \in \mathcal{J}_l^* \times \mathcal{J}_{k-l}$. But this contradicts the recursive definition of $\mathcal{J}_k^*$, establishing that the original sets must be disjoint and hence confirming that the earlier expression for ϕ_i is correct for all i by induction.

Now consider the case where the hazard function $\{\alpha_i\}_{i=1}^{\infty}$ is non-increasing for the first h periods. Using the earlier definition this implies $\mathfrak{a}_i \geq 0$ for $i = 1, \ldots, h+1$. Then the expression for ϕ_i confirms that $\phi_i \geq 0$ for these i values. Since $\varphi_i = -\phi_{i+1}$, the claim is proved.

ii. This proved by contradiction similarly to (b)ii.

iii. This follows from part (b)i. above.

A.3 Proof of Proposition 2

Let $\varsigma(z) \equiv \sum_{i=0}^{\infty} \varsigma_i z^i$ denote the z-transform of the sequence of survival probabilities $\{\varsigma_i\}_{i=0}^{\infty}$ defined in (6). That equation implies $\varsigma_i = (1-\alpha_i)\varsigma_{i-1}$, which together with Assumption 1, means that there exists a $0 < \underline{\alpha} < 1$ such that $\varsigma_i \leq (1-\underline{\alpha})\varsigma_{i-1}$. Consequently, the function $\varsigma(z)$ has certain useful properties, as stated in the following result:

Lemma 1 *Suppose $\varsigma(z) \equiv \sum_{i=0}^{\infty} \varsigma_i z^i$ is a power series with coefficients satisfying $\varsigma_0 > 0$ and $\varsigma_{i+1} \leq (1-\mathfrak{d})\varsigma_i$ for all $i \geq 0$ for some $0 < \mathfrak{d} \leq 1$. Then there exists a $\mathfrak{x} > 1$ such that:*

(i) The power series $\varsigma(z)$ is convergent and analytic on the closed disc $\mathfrak{D}_{\mathfrak{x}} \equiv \left\{z \in \mathbb{C} \mid |z| \leq \mathfrak{x}\right\}$ centred at the origin with radius $\mathfrak{x}$.

(ii) The equation $\varsigma(z) = 0$ has no roots in the set $\mathfrak{D}_{\mathfrak{x}}$.

Proof. See appendix A.1 ■

Hence by setting $\mathfrak{d} \equiv \underline{\alpha}$, there must exist a $\mathfrak{x} > 1$ such that $\varsigma(z)$ is convergent, analytic and has no zeros on $\mathfrak{D}_{\mathfrak{x}}$.

(i) Define the function $\phi(z) = 1/\varsigma(z)$, which is well-defined and analytic because of the previous result, and thus $\phi(z)$ can be written as a convergent power series for all $z \in \mathfrak{D}_{\mathfrak{x}}$. It is always the case that $\varsigma_0 = 1$, so $\phi(0) = 1$, and hence the power series can be expressed as $\phi(z) = 1 - \sum_{i=1}^{\infty} \phi_i z^i$ in terms of a sequence of coefficients $\{\phi_i\}_{i=1}^{\infty}$. Since $\phi(z)\varsigma(z) = 1$, by equating coefficients of powers of z it follows that $\phi_i = \varsigma_i - \sum_{j=1}^{i-1} \varsigma_j \phi_{i-j}$, and by making the definitions $\alpha \equiv 1 - \phi_1$ and $\varphi_i \equiv -\phi_{i+1}$, the survival function recursion recursion (8) is obtained with $n = \infty$. This is equivalent to the hazard function recursion (7) when $n = \infty$.

(ii) Since $\phi(z)$ is analytic on $\mathfrak{D}_{\mathfrak{x}}$ and $\mathfrak{x} > 1$ it must be the case that $\alpha + \sum_{i=1}^{\infty} \varphi_i < \infty$, and hence $\varphi_i \to 0$ as $i \to \infty$.

(iii) This is straightforward to prove by induction using (7).

A.4 Proof of Proposition 3

Assume $\varphi \neq 0$, since the case $\varphi = 0$ trivially requires just $0 < \alpha < 1$ for Assumption 1 to hold. Let $\phi(z) \equiv 1 - (1 - \alpha)z + \varphi z^2 = (1 - \zeta_1 z)(1 - \zeta_2 z)$ be the characteristic polynomial for the survival function recursion, where ζ_1 and ζ_2 denote the reciprocals of the roots of $\phi(z) = 0$. These satisfy $1 - \alpha = \zeta_1 + \zeta_2$ and $\varphi = \zeta_1 \zeta_2$.

Consider first the case where $\varphi > 0$. If α_i satisfies $0 < \alpha_i \leq \bar{\alpha}$ for some $0 < \bar{\alpha} < 1$ then (7) implies $0 < \alpha_{i+1} \leq \alpha + \varphi/(1 - \bar{\alpha})$. It follows that a sufficient condition for Assumption 1 to be satisfied is $\bar{\alpha} \geq \alpha + \varphi/(1 - \bar{\alpha})$ for some $0 < \bar{\alpha} < 1$. This is equivalent to $\phi\left((1 - \bar{\alpha})^{-1}\right) \leq 0$, which requires that the quadratic equation $\phi(z) = 0$ has real roots. The condition for this is $\varphi \leq \frac{1}{4}(1 - \alpha)^2$, which is one of the restrictions in (9).

If ζ_1 and ζ_2 are real then when $0 < \alpha < 1$ and $\varphi > 0$ it follows that $\zeta_1 > 0$ and $\zeta_2 > 0$. Without loss of generality, let ζ_1 be the largest reciprocal root. Then it must be the case that $0 < \zeta_1 < 1$, and hence there is a $\bar{\alpha}$ such that $0 < \bar{\alpha} < 1$ and $\phi\left((1 - \bar{\alpha})^{-1}\right) \leq 0$, guaranteeing that Assumption 1 holds.

Now consider the case $\varphi < 0$, where both roots are automatically real numbers. Here the two roots must have opposite signs, and without loss of generality, suppose $\zeta_1 > 0$ and $\zeta_2 < 0$. Since $0 < \alpha < 1$, it must be the case that $\zeta_1 > -\zeta_2$. Suppose that (9) holds. Then it follows that $\phi(z) > 0$ for all $0 \leq z \leq 1$, and because $\phi\left(\zeta_1^{-1}\right) = 0$, it must be the case that $\zeta_1 < 1$.

Since ζ_1 and ζ_2 are distinct numbers, the survival function sequence $\{\varsigma_i\}_{i=0}^{\infty}$ can be expressed as $\varsigma_i = \mathfrak{c}_1\zeta_1^i + \mathfrak{c}_2\zeta_2^i$, where $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are real numbers, and in the following alternative forms:

$$\varsigma_i = \mathfrak{c}_1\zeta_1^i \left\{1 + \frac{\mathfrak{c}_2}{\mathfrak{c}_1}\left(\frac{\zeta_2}{\zeta_1}\right)^i\right\} \quad , \quad \varsigma_i - \varsigma_{i+1} = \mathfrak{c}_1(1-\zeta_1)\zeta_1^i\left\{1 + \frac{\mathfrak{c}_2(1-\zeta_2)}{\mathfrak{c}_1(1-\zeta_1)}\left(\frac{\zeta_2}{\zeta_1}\right)^i\right\} \tag{A.3}$$

From (7) it is seen that the first two terms of the hazard function are always $\alpha_1 = \alpha$ and $\alpha_2 = \alpha + \varphi/(1-\alpha)$. Since $0 < \alpha < 1$ and $\varphi < 0$, the second term is well defined if and only if $\varphi > -\alpha(1-\alpha)$ holds. Hence given (9), the survival function satisfies $0 \leq \varsigma_2 < \varsigma_1 < \varsigma_0 = 1$. By using (A.3), $\varsigma_0 - \varsigma_1 = \mathfrak{c}_1(1-\zeta_1) + \mathfrak{c}_2(1-\zeta_2) > 0$ and $\varsigma_1 - \varsigma_2 = \mathfrak{c}_1(1-\zeta_1)\zeta_1 + \mathfrak{c}_2(1-\zeta_2)\zeta_2 > 0$. It follows from these two inequalities that $\mathfrak{c}_1 > 0$ and $\mathfrak{c}_1(1-\zeta_1) > 0$.

Because $\zeta_1 > -\zeta_2$, the terms $(\mathfrak{c}_2/\mathfrak{c}_1)(\zeta_2/\zeta_1)^i$ and $(\mathfrak{c}_2(1-\zeta_2)/\mathfrak{c}_1(1-\zeta_1))(\zeta_2/\zeta_1)^i$ must alternate in sign and decline in absolute value as i increases. Because $\mathfrak{c}_1$, ζ_1 and $(1-\zeta_1)$ are positive, the inequalities $0 \leq \varsigma_2 < \varsigma_1 < \varsigma_0$ imply $0 \leq \varsigma_i < \varsigma_{i-1}$ for all $i \geq 1$, guaranteeing that Assumption 1 holds.

To see the converse, suppose Assumption 1 is true. As the first two terms of the hazard function are well defined, the part of (9) for $\varphi < 0$ must hold. And since the survival function sequence cannot feature oscillations, this requires that the roots of $\phi(z) = 0$ are real numbers, implying the second part of (9) for $\varphi > 0$.

A.5 Proof of Proposition 4

Let $\mathcal{T}(\cdot)$ be the mapping from a distribution at time t to the distribution at time $t+1$ implied by (15), that is, $\{\theta_{i,t+1}\}_{i=0}^{\infty} = \mathcal{T}\left(\{\theta_{it}\}_{i=0}^{\infty}\right)$. Define $m \equiv \min\left\{i \mid \alpha_{i+1} = 1\right\}$ to be the maximum duration of price stickiness implied by the hazard function, noting that the case of $m = \infty$ is possible. Attention can be restricted to the set $\mathcal{S}_m \equiv \left\{\{\theta_i\}_{i=0}^m \,\middle|\, 0 \leq \theta_i \leq 1 \ ,\ \sum_{i=0}^m \theta_i = 1\right\}$ of probability distributions for the duration of price stickiness up to a maximum of m periods. Let $\mathcal{T}_m(\cdot)$ be the corresponding mapping from $\mathcal{S}_m$ to $\mathcal{S}_m$ defined by equation (15).

(i) Under Assumption 1, Lemma 1 can be applied to the z-transform $\varsigma(z) \equiv \sum_{i=0}^{\infty} \varsigma_i z^i$ of the survival probabilities $\{\varsigma_i\}_{i=0}^{\infty}$. It follows that $\sum_{i=0}^{\infty} \varsigma_i$ is finite. Define a distribution $\{\theta_i\}_{i=0}^{\infty}$ using $\theta_i \equiv \varsigma_i / \sum_{j=0}^{\infty} \varsigma_j$. This is a well-defined probability distribution, and equations (6) and (15) imply that it is a fixed point of $\mathcal{T}_m(\cdot)$.

Define $\mathcal{V}_m \equiv \left\{\{\mathsf{v}_i\}_{i=0}^m \,\middle|\, \{\mathsf{v}_i\}_{i=0}^m = \{\vartheta_i\}_{i=0}^m - \{\theta_i\}_{i=0}^m \ ,\ \{\vartheta_i\}_{i=0}^m \in \mathcal{S}_m\right\}$ to be the set of possible deviations from the first $m+1$ terms of the stationary distribution $\{\theta_i\}_{i=0}^{\infty}$. It is straightforward to see that $\mathcal{V}_m$ is a subset of the linear space $\mathcal{W}_m \equiv \left\{\{\mathsf{v}_i\}_{i=0}^m \,\middle|\, \sum_{i=0}^m \mathsf{v}_i = 0\right\}$. The definition of $\mathcal{T}_m(\cdot)$ in (15) shows that it can also be seen as a linear transformation from $\mathcal{W}_m$ to $\mathcal{W}_m$.

An eigenvalue of $\mathcal{T}_m : \mathcal{W}_m \to \mathcal{W}_m$ is a scalar $\zeta \in \mathbb{C}$ such there is a non-zero sequence $\{\mathsf{v}_i\}_{i=0}^m \in \mathcal{W}_m$ and $\mathcal{T}_m\left(\{\mathsf{v}_i\}_{i=0}^m\right) = \zeta\{\mathsf{v}_i\}_{i=0}^m$. This sequence is referred to as an eigenvector. Equation (15) shows that any eigenvalue and eigenvector pair must satisfy $\sum_{i=1}^{m+1} \alpha_i \mathsf{v}_{i-1} = \zeta \mathsf{v}_0$ and $(1-\alpha_i)\mathsf{v}_{i-1} = \zeta \mathsf{v}_i$ for all $i = 1, \ldots, m$. As the definition of the maximum duration of price stickiness m guarantees that $\alpha_i < 1$ for all $i \leq m$, it must be the case that $\zeta \neq 0$, otherwise the sequence $\{\mathsf{v}_i\}_{i=0}^{m}$ could not be an eigenvector. It follows that $\mathsf{v}_i = \varsigma_i \mathsf{v}_0/\zeta^i$ and hence $\sum_{i=0}^{m} \varsigma_i/\zeta^i = 0$, or equivalently $\varsigma\left(\zeta^{-1}\right) = 0$. Because $\varsigma(z)$ has no zeros in $\mathfrak{D}_{\mathfrak{x}}$ and $\mathfrak{x} > 1$, all the eigenvalues of $\mathcal{T}_m$ lie strictly inside the unit circle. Therefore, the stationary distribution is unique and the economy must converge to it starting from any initial conditions.

(ii) Define $\phi(z) \equiv 1 - (1-\alpha)z + \sum_{i=1}^{n} \varphi_i z^{i+1}$ to be the characteristic polynomial for the survival function recursion in (8). The functions $\phi(z)$ is then related to $\varsigma(z)$ by $\phi(z)\varsigma(z) = 1$. The definition of the stationary distribution implies $\theta_i = \varsigma_i/\varsigma(1)$, but then $\varsigma(1) = 1/\phi(1)$ and $\phi(1) = \alpha + \sum_{j=1}^{n} \varphi_j$, thus confirming the first part of (16).

The unconditional expectation of the probability of price adjustment is $\alpha^e \equiv \sum_{i=1}^{\infty} \alpha_i \theta_{i-1}$. The previous result combined with (6) implies $\alpha_i = 1-(\theta_i/\theta_{i-1})$, and so $\alpha^e = \sum_{i=0}^{\infty}(\theta_i - \theta_{i+1}) = \theta_0$, and since $\theta_0 = (\alpha + \sum_{i=1}^{n} \varphi_i)\,\varsigma_0$, the second part of (16) is shown to be correct.

The unconditional expectation of the duration of price stickiness is $\mathcal{D}^e \equiv \sum_{i=1}^{\infty} i\theta_{i-1}$. Let $\theta(z) \equiv \sum_{i=0}^{\infty} \theta_i z^i$ be the z-transform of the sequence $\{\theta_i\}_{i=0}^{\infty}$, which is analytic on $\mathfrak{D}_{\mathfrak{x}}$ because of the properties of $\varsigma(z)$. Therefore the power series can be differentiated term-by-term to yield $\theta'(z) = \sum_{i=0}^{\infty} i\theta_i z^{i-1}$. It is then seen that $\mathcal{D}^e = \theta'(1) + \theta(1)$. Since $\theta(z) = \phi(1)/\phi(z)$, it follows that $\theta'(z) = -\phi(1)\phi'(z)/(\phi(z))^2$ and hence $\mathcal{D}^e = (\phi(1) - \phi'(1))/\phi(1)$. The definition of $\phi(z)$ yields $\phi'(z) = -(1-\alpha) + \sum_{i=1}^{n}(i+1)\varphi_i z^i$, proving the third part of (16).

A.6 Proof of Theorem 1

Define the sequence $\{\phi_i\}_{i=1}^{n+1}$ with $\phi_1 \equiv 1-\alpha$ and $\phi_i = -\varphi_{i-1}$ for $i \geq 2$, and set $\phi(z) \equiv 1 - \sum_{i=1}^{n+1} \phi_i z^i$. Then let $\varsigma(z) \equiv \sum_{i=0}^{\infty} \varsigma_i z^i$ and $\theta(z) \equiv \sum_{i=0}^{\infty} \theta_i z^i$ be the z-transforms of the survival function and the duration distribution respectively. As Assumption 1 holds, Lemma 1 and equation (8) imply that $\phi(z)\varsigma(z) = 1$ for all z in the unit disc $\mathfrak{D}$. It then follows from equation (16) of Proposition 4 that $\phi(z)\theta(z) = \phi(1)$ for all $z \in \mathfrak{D}$. In terms of the lag operator $\mathbb{L}$ and the forward operator $\mathbb{F}$, the equations in (19) can be expressed as $\mathsf{R}_t = \mathbb{E}_t[(\varsigma(\beta\mathbb{F})/\varsigma(\beta))(\mathsf{P}_t + \eta_{cx}\mathsf{x}_t)]$ and $\mathsf{P}_t = \theta(\mathbb{L})\mathsf{R}_t$. The equivalent recursive versions in (20) and (22) are $\mathbb{E}_t[\phi(\beta\mathbb{F})\mathsf{R}_t] = \phi(\beta)(\mathsf{P}_t + \eta_{cx}\mathsf{x}_t)$ and $\phi(\mathbb{L})\mathsf{P}_t = \phi(1)\mathsf{R}_t$.

By substituting the recursive equation for the reset price into the recursive equation for the price level, the stochastic difference equation $\mathbb{E}_t[\Upsilon(\mathbb{L})\mathsf{P}_t] = \phi\eta_{cx}\mathsf{x}_t$ is obtained with definitions $\Upsilon(z) \equiv \phi(z)\phi(\beta z^{-1}) - \phi(1)\phi(\beta)$ and $\phi \equiv \phi(1)\phi(\beta)$. The function $\Upsilon(z)$ has the symmetry property $\Upsilon(z) = \Upsilon(\beta z^{-1})$ for all $z \neq 0$. Together with the definition of $\phi(z)$ this implies that $\Upsilon(z)$ can be written as $\Upsilon(z) \equiv \Upsilon_0 + \sum_{i=1}^{n+1} \Upsilon_i\left(z^i + \beta^i z^{-i}\right)$ with $\Upsilon_i = -\left(\phi_i - \sum_{j=1}^{n+1-i} \beta^j \phi_j \phi_{i+j}\right)$ for $i \geq 1$.

The definition of $\Upsilon(z)$ ensures that $\Upsilon(1) = 0$. It follows that there exists a function $\psi(z) \equiv 1 - \sum_{i=1}^{n} \psi_i z^i - \sum_{i=1}^{n+1} \delta_i z^{-i}$ and a constant κ such that $\Upsilon(z) = (\phi/\kappa)(1-z)\psi(z)$. This allows the stochastic difference equation involving prices and real marginal cost to be replaced by an equation $\mathbb{E}_t[\psi(\mathbb{L})\pi_t] = \kappa\eta_{cx}\mathsf{x}_t$ in terms of inflation and real marginal cost. If this equation is written out in full and $\kappa_x \equiv \kappa\eta_{cx}$ is defined then the Phillips curve given in (23) is obtained.

Expressions for the coefficients of the Phillips curve in terms of the hazard function parameters are obtained by using the relationship between the $\Upsilon(z)$ and $\psi(z)$ functions and the expression derived for Υ_i. The coefficients on lagged inflation are given by the expression $\psi_i = -(\kappa/\phi) \sum_{j=i+1}^{n+1} \phi_j \left(1 - \sum_{k=1}^{j-(i+1)} \beta^k \phi_k\right)$. Similarly, the coefficients on inflation expectations and real marginal cost are $\delta_{i+1} = (\kappa/\phi) \sum_{j=i+1}^{n+1} \beta^j \phi_j \left(1 - \sum_{k=1}^{j-(i+1)} \phi_k\right)$ and $\kappa = \phi / \left(\sum_{j=1}^{n+1} \phi_j \left(1 - \sum_{k=1}^{j-1} \beta^k \phi_k\right)\right)$ respectively. By using the definitions made at the beginning of this proof, all the coefficients can be stated in terms of α, $\{\varphi_i\}_{i=1}^n$, β and η_{cx}:

$$\psi_i = \frac{\varphi_i + \sum_{j=i+1}^{n} \varphi_j \left(1 - \beta(1-\alpha) + \sum_{k=1}^{j-1} \beta^{k+1} \varphi_k\right)}{(1-\alpha) - \sum_{j=1}^{n} \varphi_j \left(1 - \beta(1-\alpha) + \sum_{k=1}^{j-1} \beta^{k+1} \varphi_k\right)} \qquad i = 1, \ldots, n \tag{A.4a}$$

$$\delta_{i+1} = -\beta^{i+1} \frac{\varphi_i + \sum_{j=i+1}^{n} \beta^{j-i} \varphi_j \left(\alpha + \sum_{k=1}^{j-1} \varphi_k\right)}{(1-\alpha) - \sum_{j=1}^{n} \varphi_j \left(1 - \beta(1-\alpha) + \sum_{k=1}^{j-1} \beta^{k+1} \varphi_k\right)} \qquad i = 1, \ldots, n \tag{A.4b}$$

$$\delta_1 = \beta \frac{(1-\alpha) - \sum_{j=1}^{n} \beta^j \varphi_j \left(\alpha + \sum_{k=1}^{j-1} \varphi_k\right)}{(1-\alpha) - \sum_{j=1}^{n} \varphi_j \left(1 - \beta(1-\alpha) + \sum_{k=1}^{j-1} \beta^{k+1} \varphi_k\right)} \tag{A.4c}$$

$$\kappa_x = \frac{\eta_{cx} \left(\alpha + \sum_{j=1}^{n} \varphi_j\right) \left(1 - \beta(1-\alpha) + \sum_{j=1}^{n} \beta^{j+1} \varphi_j\right)}{(1-\alpha) - \sum_{j=1}^{n} \varphi_j \left(1 - \beta(1-\alpha) + \sum_{k=1}^{j-1} \beta^{k+1} \varphi_k\right)} \tag{A.4d}$$

To establish the link between the signs of the sequences $\{\varphi_i\}_{i=1}^n$ and $\{\psi_i\}_{i=1}^n$, note $\phi(z)\varsigma(z) = 1$, $0 < \beta < 1$ and Assumption 1 imply $0 < \alpha + \sum_{i=1}^{n} \varphi_i < 1$ and $0 < 1 - \beta(1-\alpha) + \sum_{i=1}^{n} \beta^{i+1} \varphi_i < 1$.

(a) This follows immediately from (A.4a) by induction.

(b) i. Consider the case where $\varphi_i \geq 0$ for all $i = 1, \ldots, n$. Since Assumption 1 requires that $0 < \alpha < 1$, the inequalities stated above can be extended to yield $0 < \alpha + \sum_{k=1}^{j-1} \varphi_k < 1$ and $0 < 1 - \beta(1-\alpha) + \sum_{k=1}^{j-1} \beta^{k+1} \varphi_k < 1$ for all $j = 1, \ldots, n+1$. Therefore, by using the expression for ψ_i in (A.4a), $\varphi_i \geq 0$ for all $i = 1, \ldots, n$ implies $\psi_i \geq 0$ for $i = 1, \ldots, n$. Similarly, $\varphi_i > 0$ for all i yields $\psi_i > 0$ for all i.

ii. The inequalities derived in part (b)i. also hold. The result follows immediately from (A.4a).

iii. This is the contrapositive of the claim proved in (c)i. below.

(c) i. Now suppose $\varphi_i \leq 0$ for all $i = 1, \ldots, n$. Using similar reasoning to part (b)i. above, the inequalities derived there must hold for all $j = 1, \ldots, n+1$. Hence from (A.4a), $\varphi_i \leq 0$ for all $i = 1, \ldots, n$ implies $\psi_i \leq 0$ for $i = 1, \ldots, n$. In addition, $\varphi_i < 0$ for all i leads to $\psi_i < 0$ for all i.

ii. The inequalities derived in part (b)i. also hold, and the result follows from (A.4a).

iii. This is the contrapositive of part (b)i. above, which has already been proved.

References

Altissimo, F., Bilke, L., Levin, A., Mathä, T. and Mojon, B. (2006), 'Sectoral and aggregate inflation dynamics in the euro area', Journal of the European Economic Association, 4(2–3), pp. 585–593.

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.

Baumgartner, J., Glatzer, E., Rumler, F. and Stiglbauer, A. (2005), 'How frequently do consumer prices change in Austria? Evidence from micro CPI data', Working paper 523, European Central Bank.

Calvo, G. A. (1983), 'Staggered prices in a utility-maximizing framework', Journal of Monetary Economics, 12, pp. 383–398.

Campbell, J. R. and Eden, B. (2005), 'Rigid prices: Evidence from U.S. scanner data', Working paper, Vanderbilt University.

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

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

Cliff, M. T. (2003), 'GMM and MINZ program libraries for MATLAB', Working paper, Virginia Tech, available for download from `mcliff.cob.vt.edu/progs.html`.

Cogley, T. and Sbordone, A. M. (2005), 'A search for a structural Phillips curve', Staff report 203, Federal Reserve Bank of New York.

Dias, D. A., Marques, C. R. and Santos Silva, J. M. C. (2005), 'Time or state dependent price setting rules? Evidence from Portuguese micro data', Working paper 511, European Central Bank.

Dittmar, R. D., Gavin, W. T. and Kydland, F. E. (2005), 'Inflation persistence and flexible prices', International Economic Review, 46(1), pp. 245–261.

Dotsey, M. (2002), 'Pitfalls in interpreting tests of backward-looking pricing in New Keynesian models', Federal Reserve Bank of Richmond Economic Quarterly, 88, pp. 37–50.

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), pp. 655–690.

Fougère, D., Le Bihan, H. and Sevestre, P. (2005), 'Heterogeneity in consumer price stickiness: A microeconometric investigation', Working paper 536, European Central Bank.

Fuhrer, J. C. (1997), 'The (un)importance of forward-looking behavior in price specifications', Journal of Money, Credit and Banking, 29, pp. 338–350.

Fuhrer, J. C. and Moore, G. R. (1995), 'Inflation persistence', Quarterly Journal of Economics, 109, pp. 197–230.

Gadzinski, G. and Orlandi, F. (2004), 'Inflation persistence in the European Union, the Euro Area, and the United States', Working paper 414, European Central Bank.

Galí, J. and Gertler, M. (1999), 'Inflation dynamics: A structural econometric analysis', Journal of Monetary Economics, 44, pp. 195–222.

Galí, J., Gertler, M. and López-Salido, J. D. (2001), 'European inflation dynamics', European Economic Review, 45, pp. 1237–1270.

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

Goodfriend, M. and King, R. G. (1997), 'The New Neoclassical synthesis and the role of monetary policy', NBER Macroeconomics Annual, 12, pp. 231–283.

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.

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

Guerrieri, L. (2002), 'The inflation persistence of staggered contracts', International finance discussion paper 734, Federal Reserve Board.

Heckman, J. and Singer, B. (1984), 'Econometric duration analysis', Journal of Econometrics, 24, pp. 63–132.

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

Kiefer, N. (1988), 'Economic duration data and hazard functions', Journal of Economic Literature, 26, pp. 646–679.

Kurmann, A. (2004), 'Maximum likelihood estimation of dynamic stochastic theories with an application to New Keynesian pricing', Working paper 04–21, Université du Québec à Montréal.

Lindé, J. (2005), 'Estimating New-Keynesian Phillips curves: A full information maximum likelihood approach', Journal of Monetary Economics, 52(6), pp. 1135–1149.

Mankiw, N. G. (2001), 'The inexorable and mysterious trade-off between inflation and unemployment', Economic Journal, 111, pp. C45–61.

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

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

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

Nakamura, E. and Steinsson, J. (2007), 'Five facts about prices: A reevaluation of menu cost models', Working paper, Harvard University.

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

Roberts, J. M. (1997), 'Is inflation sticky?', Journal of Monetary Economics, 39, pp. 173–196.

Roberts, J. M. (2005), 'How well does the New Keynesian sticky-price model fit the data?', Contributions to Macroeconomics, 5(1), p. Article 10.

Rudd, J. and Whelan, K. (2005), 'New tests of the New-Keynesian Phillips curve', Journal of Monetary Economics, 52(6), pp. 1167–1181.

Sbordone, A. (2002), 'Prices and unit labor costs: A new test of price stickiness', Journal of Monetary Economics, 49, pp. 235–256.

Sheedy, K. D. (2007), 'Robustly optimal monetary policy', Working paper, London School of Economics.

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), pp. 1123–1175.

Taylor, J. B. (1980), 'Aggregate dynamics and staggered contracts', Journal of Political Economy, 88, pp. 1–24.

Whelan, K. (2007), 'Staggered price contracts and inflation persistence: Some general results', International Economic Review, 48(1), pp. 111–145.

Wolman, A. L. (1999), 'Sticky prices, marginal cost, and the behavior of inflation', Federal Reserve Bank of Richmond Economic Quarterly, 85(4), pp. 29–48.

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

Woodford, M. (2007), 'Interpreting inflation persistence: Comments on the conference on 'Quantitative Evidence on Price Determination'', Journal of Money, Credit and Banking, 39(1), pp. 203–210.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

836	Andrew E. Clark Ed Diener Yannis Georgellis Richard E. Lucas	Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis
835	Jean-Baptiste Michau	Creative Destruction with On-the-Job Search
834	Nikolaus Wolf	Scylla and Charybodis: the European Economy and Poland's Adherence to Gold, 1928-1936
833	Sharon Belenzon Tomer Berkovitz	Innovation in Business Groups
832	Espen R. Moen Åsa Rosén	Incentives in Competitive Search Equilibrium
831	Eran Yashiv	U.S. Labor Market Dynamics Revisited
830	Laura Alfaro Andrew Charlton	Growth and the Quality of Foreign Direct Investment: Is All FDI Equal?
829	Richard Layard David Clark Martin Knapp Guy Mayraz	Cost-Benefit Analysis of Psychological Therapy
828	Emma Tominey	Maternal Smoking During Pregnancy and Early Child Outcomes
827	Christos Genakos Tommaso Valletti	Testing the "Waterbed" Effect in Mobile Telephony
826	Luis Garicano Paul Heaton	Information Technology, Organization, and Productivity in the Public Sector: Evidence from Police Departments
825	Laura Alfaro Andrew Charlton	Intra-Industry Foreign Direct Investment
824	Miklós Koren Silvana Tenreyro	Technological Diversification
823	Régis Barnichon	The Shimer Puzzle and the Correct Identification of Productivity Shocks
822	Carlos Thomas	Search Frictions, Real Rigidities and Inflation Dynamics

821	Giulia Faggio Kjell G. Salvanes John Van Reenen	The Evolution of Inequality in Productivity and Wages: Panel Data Evidence
820	Luis Araujo Emanuel Ornelas	Trust-Based Trade
819	Régis Barnichon	Productivity, Aggregate Demand and Unemployment Fluctuations
818	Rachel Griffith Sokbae Lee John Van Reenen	Is Distance Dying at Last? Falling Home Bias in Fixed Effects Models of Patent Citations
817	Tsutomu Miyagawa YoungGak Kim	Measuring Organization Capital in Japan: An Empirical Assessment Using Firm-Level Data
816	Andreas Georgiadis Alan Manning	Spend It Like Beckham? Inequality and Redistribution in the UK, 1983-2004
815	Guy Michaels Xiaojia Zhi	Freedom Fries
814	Ralph Ossa	Trade Liberalization, Outsourcing, and Firm Productivity
813	Karsten Kohn Alexander C. Lembcke	Wage Distributions by Bargaining Regime: Linked Employer-Employee Data Evidence from Germany
812	Yannis M. Ioannides Henry G. Overman Esteban Rossi-Hansberg Kurt Schmidheiny	The Effect of Information and Communication Technologies on Urban Structure
811	Guy Michaels	The Division of Labor, Coordination, and the Demand for Information Processing
810	Maia Güell Jose V. Rodriguez Mora Chris Telmer	Intergenerational Mobility and the Informative Content of Surnames
809	Stephen J. Redding Daniel M. Sturm Nikolaus Wolf	History and Industry Location: Evidence from German Airports
808	Anna Lipinska	The Maastricht Convergence Criteria and Optimal Monetary Policy for the EMU Accession Countries
807	Eran Yashiv	The Beveridge Curve

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk