Changes in U.S. Inflation Persistence

Kyu Ho Kang∗ Chang-Jin Kim†
James Morley‡

∗Washington University in St. Louis, khkang@wustl.edu
†Korea University and University of Washington, changjin@u.washington.edu
‡Washington University in St. Louis, morley@wustl.edu
Changes in U.S. Inflation Persistence*

Kyu Ho Kang, Chang-Jin Kim, and James Morley

Abstract

We investigate the existence and timing of changes in U.S. inflation persistence. To do so, we develop an unobserved components model of inflation with Markov-switching parameters and we measure persistence using impulse response functions based on the model. An important feature of our model is its allowance for multiple regime shifts in parameters related to the size and propagation of shocks. Inflation persistence depends on the configuration of these parameters, although it need not change even if the parameters change. Using the GDP deflator for the sample period of 1959-2006, we find that U.S. inflation underwent two sudden permanent regime shifts, both of which corresponded to changes in persistence. The first regime shift occurred around the collapse of the Bretton Woods system at the beginning of the 1970’s and produced an increase in inflation persistence, while the second regime shift occurred immediately after the Volcker disinflation in the early 1980’s and produced a decrease in inflation persistence. Meanwhile, consistent with the New Keynesian Phillips Curve, the gap between inflation and its long-run trend displayed little or no persistence throughout the entire sample period.

*We thank the editor, an anonymous referee, Jeremy Piger, and participants at the ESRC Seminar Series on Nonlinear Economics and Finance Research Community 2007 for helpful comments and suggestions. Kim acknowledges financial support from the Korea Research Foundation Grant (KRF-2008-342-B00006) and from the Bryan C. Cressey Professorship at the University of Washington. Morley acknowledges support from the Weidenbaum Center on the Economy, Government, and Public Policy. The usual disclaimers apply.
1. Introduction

Theoretically speaking, the persistence of inflation is a major determinant of the economic costs of a disinflation.\footnote{Fuhrer and Moore (1995) examine theoretical and empirical issues surrounding inflation persistence and the costs of a disinflation. They show that a staggered wage contract model in which agents care about relative real wages (also see Buiter and Jewett, 1981) can explain persistent inflation and implies a much larger sacrifice ratio than models with less persistent inflation (also see Fuhrer, 1995). Meanwhile, using a long history of data, Bordo and Haubrich (2004) find that changes in inflation persistence can help explain changes in ability of the yield curve to predict the U.S. business cycle, with the implication being that higher inflation persistence corresponds to larger declines in aggregate output following contractionary monetary policy.} Furthermore, Cogley and Sargent (2001) argue that changes in inflation persistence can explain the policy mistakes that lead to high inflation in the first place. In general, inflation persistence is important for determining the relevance of the widely-used New Keynesian Phillips Curve based on price stickiness.\footnote{Mankiw (2001) takes a critical view of the New Keynesian Phillips Curve given that it cannot generate persistent inflation without the unsatisfactory assumptions of serially correlated exogenous shocks or backward-looking expectations. However, Cogley and Sbordone (2008) argue that the New Keynesian Phillips Curve is only designed to explain the degree of persistence in the gap of inflation from its long-run trend, where the trend changes over time due to shifts in monetary policy (i.e., changes in policy rules or in the long-run inflation target).} Despite (or perhaps because of) its importance for macroeconomics, the degree of inflation persistence and whether it has changed remains a controversial and unresolved empirical issue. Meanwhile, even when it is argued that persistence has changed, there are vastly different findings on the timing and nature of such changes, often due to restrictive econometric assumptions. Thus, the basic goal of this paper is to investigate the existence and the nature of changes in U.S. inflation persistence using a time series modeling framework that allows for a wide range of possibilities.

In recent work, Stock and Watson (2007) and Pivetta and Reis (2007) argue that U.S. inflation persistence has remained unchanged for many decades. In particular, Stock and Watson (2007) measure the persistence as the largest autoregressive root in the inflation rate, for which the 90% confidence interval is found to contain a unit root for both the subsamples of 1970-1983 and 1984-2004. Similarly, by estimating a Bayesian time-varying parameter model, Pivetta and Reis (2007) find that the largest autoregressive root was close to one and
essentially constant over the entire sample period of 1947-2001. However, these findings stand in contrast to earlier work that also considers the largest autoregressive root. For example, Kim (2000), using data from 1948-1994, presents evidence that the U.S. inflation rate switched from being stationary to a unit root process in 1973, while Leybourne, Kim, Smith and Newbold (2003), using data from 1959-2000, present evidence that the U.S. inflation switched from being a unit root process to being stationary in 1982. Meanwhile, even if these earlier studies are correct about the existence of structural change in inflation persistence, they have the crucial limitation of assuming just a single unknown break date. Indeed, their different results are directly suggestive of the possibility of multiple regime shifts in the inflation process.

In this paper, we develop an unobserved components model of inflation with Markov-switching parameters that allows for multiple regime shifts and places no restriction that regime shifts need to be between stationarity and a unit root. Importantly, the regime shifts can have permanent or transitory effects on the parameters related to the size and propagation of shocks, thus accommodating heteroskedasticity and allowing persistence to change across regimes. By measuring persistence using implied impulse response functions, we find that

---

3 In addition to focusing on the largest autoregressive root, Pivetta and Reis (2007) also consider two other measures of persistence: the sum of the autoregressive parameters in their time-varying parameter model and the half-life response of inflation to a shock according to their model. They find high and stable persistence for all three measures.

4 Despite similar technical approaches, the difference in results for the timing and nature of structural change lies in the fact that Kim’s (2000) test is taken against the single alternative of stationarity changing to a unit root, while Leybourne, Kim, Smith and Newbold (2003) analyze the inflation dynamics without prior knowledge of the direction of the change.

5 There are a few studies that directly allow for multiple regime shifts in inflation persistence. Evans and Wachtel (1993) develop a Markov-switching model of inflation with two regimes, a stationary regime and a unit root regime, and find that U.S. inflation switched from stationarity to a unit root regime in the early-to-mid-1970s and returned to the stationarity in the mid-1980s. More recently, Murray, Nikolsko-Rzhevskyy, and Papell (2009) consider a similar model and find similar timing for shifts between a stationary regime and an explosive regime. Meanwhile, Burdekin and Siklos (1999) apply the Perron and Vogelsang (1992) procedure to sequentially test for multiple structural breaks at unknown breakpoints to annual CPI inflation and find one significant break in U.S. inflation persistence in 1979 given a sample period of 1946-1993. However, the test does not account for the possibility of heteroskedasticity, while the existence of heteroskedasticity for inflation has been verified in a number of studies, including Evans and Wachtel (1993) and Kim (1993).
there were two permanent changes in the persistence of U.S. inflation over the sample period of 1959-2006. The changes are estimated to have occurred in 1970Q3 and 1984Q4 and were large in magnitude, with inflation persistence increasing in with the collapse of the Bretton Woods system in the 1970s and decreasing after the Volcker disinflation in the mid-1980s. Meanwhile, we find that the gap between inflation and its long-run trend displayed little or no persistence throughout the entire sample period, consistent with the New Keynesian Phillips Curve.

Our results can be related to other findings in the literature. First, in terms of U.S. inflation persistence, Evans and Wachtel (1993) estimate the same “low-high-low” pattern for the sample period of 1955-1991, although their regime-switching model assumes recurrent switching between a stationary regime and a unit root regime, while we find permanent switches between three unit root regimes with different degrees of persistence. Cogley and Sargent (2001, 2005) also find the same basic pattern for inflation persistence over the post-World War II sample period, although their Bayesian time-varying parameter model assumes that changes in persistence were more gradual than what can be found with a Markov-switching model. Stock and Watson (2007) employ an unobserved components model of inflation that is similar to the one considered here, although they assume stochastic volatility instead of Markov switching and a more restrictive specification for the gap between inflation and its trend (they assume a priori that the inflation gap is serially uncorrelated). Regardless, they find broadly consistent results in terms of the relative importance of permanent shocks, even though their focus on the largest autoregressive root leads them to conclude that there were no changes in inflation persistence. Second, in terms of the inflation gap, our finding of low persistence is broadly consistent with the international evidence, including for the United States, reported in Levin and Piger (2006), who measure the inflation gap by allowing for a one-time structural break in the level of inflation for the relatively short sample period of 1984-2004. Our results for the inflation gap are also consistent with the findings in Cogley, Primiceri, and Sargent (2007) and Cogley and Sbordone (2008), who consider Bayesian time-varying parameter models and the Beveridge-Nelson decomposition to measure
long-run trend inflation, although the estimated persistence in our inflation gap is even lower than theirs, potentially providing even more support for the forward-looking version of the New Keynesian Phillips Curve.\(^6\)

The approach taken in this paper contributes to the existing literature in two ways. First, we provide arguments for measuring inflation persistence using the impulse response function for inflation given a forecast error rather than the largest autoregressive root or other measures that are more suitable for comparing stationary processes. Succinctly, the long-run response is able to discriminate between a unit root process that is largely subject to permanent variation and one that is largely subject to transitory variation, whereas the largest autoregressive root considered in many previous studies only answers the simpler question of whether or not the process has any permanent variation. Second, the time series model considered here is flexible in the sense that it can accommodate the various findings in previous studies mentioned above, yet it is informative in the sense that it delivers a clear finding of two permanent changes in inflation persistence that are closely related to major changes in the U.S. monetary policy environment. Importantly, the model explicitly allows for sudden changes in persistence rather than smooth changes that are assumed \textit{a priori} in time-varying parameter models (see, for example, Pivetta and Reis, 2007). This allowance for sudden changes is particularly relevant for inflation data given the idea put forth by many researchers (e.g., Cogley and Sargent, 2001, Cogley and Sbordone, 2008, and Benati, 2008) and confirmed in this paper that inflation persistence is interwoven with shifts in monetary policy.

The remainder of this paper is organized as follows. In Section 2, we discuss measures of persistence. Section 3 presents the unobserved components model used to detect changes in U.S. inflation dynamics. Section 4 reports the

\(^6\) Benati (2008) confirms Levin and Piger’s (2006) international evidence of low inflation persistence during monetary policy regimes with a well-defined nominal anchor, but takes the view that low persistence does not rescue the New Keynesian Phillips Curve given instabilities in structural parameters across policy regimes for estimated DSGE models with sticky prices. Meanwhile, his link between inflation persistence and monetary policy regimes is consistent with our findings in terms of a close link between the timing and direction of changes in inflation persistence and the collapse of the Bretton Woods system in the early 1970s and the Volcker disinflation in the mid-1980s.
estimation results for the model, including impulse response functions and inferences about the timing of regime shifts. Section 5 concludes.

2. Measures of Persistence

A question of crucial importance is how to measure the persistence of a time series process. Following Andrews and Chen (1994), we consider impulse response functions (IRFs), which measure the dynamic effects of shocks on a time series process over different horizons. At a basic level, an IRF captures persistence given that the response coefficients will eventually die out to zero for a stationary process, but will not do so for a unit root process. Of course, this distinction alone does not explain why an IRF is a preferable measure of persistence than, say, the largest autoregressive root, which is less than one (in absolute value) for a stationary process and equal to one for a unit root process.

The benefit of IRFs is that they can capture the degree of persistence, even when considering a unit root process. Given that U.S. inflation is typically assumed to have a unit root, including by researchers with different views about whether its persistence has actually changed (see, for example, Stock and Watson, 2007, and Cogley, Primiceri, and Sargent, 2007), it makes sense to consider a measure of inflation persistence that can vary given the presence of a unit root. The largest autoregressive root or the sum of the autoregressive coefficients for inflation are always equal to one if it follows a unit root process, regardless of whether there are changes in the relative permanent and transitory variation in inflation. By contrast, the long-run response of inflation to a forecast error implied by an IRF does vary with the relative importance of permanent or transitory shocks. If most variation in inflation is permanent, then the long-run response to a one-unit forecast error will be close to one, while it will be close to zero if most of the variation in inflation is transitory. Meanwhile, even if inflation follows a stationary process, we can still use IRFs to measure the degree of persistence. The long-run response will always be zero, but the short-run IRFs can be different, as measured by, say, the half-life of a response to a forecast error.

Using the long-run response of inflation to a one-unit forecast error to
measure inflation persistence is equivalent to considering the normalized spectral density at frequency zero for the first differences of inflation. Meanwhile, if the first differences of inflation can be modeled by an autoregressive (AR) process, then the spectral density at frequency zero would be a monotone transformation of the sum of the autoregressive coefficients (see Andrews and Chen, 1994). Thus, it can make sense to look at the sum of autoregressive coefficients to measure persistence. However, there are two subtle issues that need to be addressed. First, the sum of the autoregressive coefficients for an AR process captures persistence for the accumulation of the process. Thus, if inflation were stationary, the sum of its autoregressive coefficients would capture the long-run persistence of the log price level, not inflation per se. Second, if inflation follows a unobserved components process with a finite-order AR transitory component, then its first differences will have moving average (MA) terms, meaning that the sum of the autoregressive coefficients for the first differences will not directly relate to the spectral density at frequency zero. Thus, for inflation, we measure persistence by considering the cumulative impulse response function implied by an ARMA model of its first differences, rather than considering its largest autoregressive root or the sum of the autoregressive coefficients for its first differences.

3. Model Specification

For our analysis, we consider an unobserved components (UC) model of inflation with Markov-switching parameters. The basic UC structure is similar to that in Ball and Cecchetti (1990), Nason (2006), and Stock and Watson (2007), with the inflation rate \( \pi_t \) assumed to be the sum of a random walk \( \tau_t \) and a stationary ARMA(p, q) process \( c_t \):

\[
\pi_t = \tau_t + c_t \tag{1}
\]

\[
\tau_t = \mu_S + \tau_{t-1} + \eta_t \tag{2}
\]

\[
\phi_S(L)c_t = \varphi_S(L)\epsilon_t \tag{3}
\]
\[
\begin{bmatrix}
\eta_t \\
\varepsilon_t
\end{bmatrix}
\sim iidN(0, Q_{S_t}),
\] (4)

where

\[
\phi_{S_t}(L) = 1 - \phi_{1,S_t} L - \phi_{2,S_t} L^2 - \cdots - \phi_{p,S_t} L^p
\] (5)

\[
\varphi_{S_t}(L) = 1 - \varphi_{1,S_t} L - \varphi_{2,S_t} L^2 - \cdots - \varphi_{q,S_t} L^q
\] (6)

\[
Q_{S_t} = \begin{bmatrix}
\sigma_{\eta,S_t}^2 & \rho_{\eta,S_t} \sigma_{\varepsilon,S_t} \\
\rho_{\eta,S_t} \sigma_{\varepsilon,S_t} & \sigma_{\varepsilon,S_t}^2
\end{bmatrix},
\] (7)

Ball and Cecchetti (1990), Nason (2006), and Stock and Watson (2006) all impose the restriction of zero correlation between the permanent and transitory shocks. However, following Morley, Nelson, and Zivot (2003), we assume that the two shocks can be correlated, with the correlation parameter \( \rho_{S_t} \) being identified as long as \( p \geq q + 2 \) for the ARMA(p, q) process.\(^7\) Meanwhile, for the stationarity of \( c_t \), we assume that the roots that solve the characteristic equation

\[
1 - \phi_{1,S_t} z - \phi_{2,S_t} z^2 - \cdots - \phi_{p,S_t} z^p = 0
\] for the autoregressive lag polynomial all lie outside the unit circle.

The main distinguishing feature of our model is the allowance for possible discrete regime shifts in inflation persistence by letting the model parameters in (1)-(7) depend on an \( N \)-state Markov-switching variable \( S_t \), with fixed transition probabilities:

\[
\Pr[S_t = i \mid S_{t-1} = j] = p_{ij} \text{ for } i, j \in \{1, 2, \ldots, N\}.
\] (8)

Furthermore, as discussed in the next section, we find that regimes are so persistent as to be essentially permanent. Thus, we also consider restricted versions of the Markov-switching model in which the regime changes are permanent. For notational convenience with the restricted model, we normalize

\(^7\) Our model also differs from the previous literature by the inclusion of a state-dependent drift term in the permanent component. This term prevents the size of the permanent shocks from being overestimated in the case that inflation drifts upwards or downwards in any of the regimes, but can be estimated to be zero in the case that there is no inherent drift.
the labels for the regimes based on their probable sequence of occurrence, which we determine *ex post* based on smoothed probabilities. Specifically, Regime 1 is the regime that was most likely occurring at the beginning of the sample, Regime 2 is the regime that most likely occurred following the first transition out of Regime 1, Regime 3 is the regime that most likely occurred following the (possible multiple) occurrences of Regimes 1 and 2, etc… Then, in order to consider permanent regime changes, we impose the following restrictions on transition probabilities for the Markov-switching variable:

\[
\Pr[S_t = i + 1 | S_{t-1} = i] = 1 - p_{ii} \quad \text{for} \quad i \in \{1, 2, \ldots, N-1\} \quad \text{and} \quad p_{NN} = 1 \quad (8')
\]

That is, the regimes are “terminal” (i.e., once left, they will never be returned to) and the last regime is “absorbing” (i.e., once entered, it will never be left).

Given the UC structure, the overall persistence of inflation depends on the relative size of permanent and transitory shocks, as well as the ARMA propagation of the transitory shocks. To calculate a scalar measure of persistence that reflects these various factors, we transform the unobserved components model into its reduced-form ARIMA representation and solve for the implied regime-dependent cumulative impulse response functions to obtain the expected long-run response of inflation to a forecast error in different regimes.

The ARIMA representation for the UC model is given by

\[
\phi_S(L) \Delta \pi_t = \theta_S(L) e_t, \quad (9)
\]

\[
e_t \sim iidN(0, \sigma_{e, S}^2), \quad (10)
\]

where

\[
\theta_S(L) = 1 - \theta_{1, S} L - \theta_{2, S} L^2 - \cdots - \theta_{k, S} L^k,
\]

with \( k = \max\{p, q + 1\} \). The MA parameters and the forecast error variance are complicated functions of the UC model parameters. It is important to emphasize that this model directly accommodates the possibility raised by Sims (2001) that regime shifts correspond only to heteroskedasticity rather than changes in conditional mean dynamics. In particular, it is possible that only the variance of the
forecast error changes across regimes, with the ARMA parameters remaining essentially unchanged.

4. Empirical Results

Following much of the previous literature (e.g., Pivetta and Reis, 2007, and Stock and Watson, 2007), we consider data for the U.S. GDP deflator. The sample period is 1959Q1-2006Q2 and the inflation rate is calculated as the annualized quarterly percentage change in the price index.

4.1. Determining the Number of Regimes

For Markov-switching models, formal hypothesis tests for the number of regimes are confounded by identically zero scores at the null and the presence of nuisance parameters under the alternative (see Hansen, 1992, and Garcia, 1998). Thus, we address the practical question of how many regimes to include in our model in two less formal ways. First, we consider model selection based on the Akaike Information Criterion (AIC). Second, we verify our model selection results by conducting residual diagnostic tests. Specifically, we consider whether a given model captures all of the serial correlation and heteroskedasticity in the data using the modified Ljung-Box portmanteau tests for the standardized residuals and squared standardized residuals. Looking at the results in Table 1 and focusing on versions of the UC model with recurring regime shifts, AIC selects two regimes

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>$\ln L$</th>
<th>$k$</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No changes</td>
<td>-286.125</td>
<td>6</td>
<td>-292.125</td>
</tr>
<tr>
<td>Two Recurring Regimes</td>
<td>-266.044</td>
<td>14</td>
<td>-280.044</td>
</tr>
<tr>
<td>Two Permanent Regimes</td>
<td>-280.672</td>
<td>13</td>
<td>-293.672</td>
</tr>
<tr>
<td>Three Recurring Regimes</td>
<td>-259.956</td>
<td>24</td>
<td>-283.957</td>
</tr>
<tr>
<td>Three Permanent Regimes</td>
<td>-259.150</td>
<td>20</td>
<td>-279.150</td>
</tr>
</tbody>
</table>

Notes: $\ln L$ denotes the log likelihood and $k$ is the number of model parameters. AIC denotes the Akaike Information Criterion and is calculated as $\ln L - k$, meaning that larger statistics correspond to smaller estimated Kullback-Leibler distances from the true model.
instead of three. However, when considering models with more restrictive permanent regime shifts with transition probabilities given (8’) instead of (8), AIC selects the version of the model with three permanent regimes instead of two. Furthermore, the unrestricted version of the model with three regimes fits the data in almost the identical way as the restrictive version, suggesting no need for a more complicated model. More formally, the LR test statistic for the hypothesis that the three regimes are permanent ( \( p_{13} = p_{21} = p_{32} = p_{31} = 0 \) ) is only 1.612, which is

<table>
<thead>
<tr>
<th>Lag</th>
<th>( Q^* )-statistic</th>
<th>p-value</th>
<th>( Q^* )-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standardized Residuals</td>
<td>Squared Standardized Residuals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. No changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.100</td>
<td>0.752</td>
<td>3.675</td>
<td>0.055</td>
</tr>
<tr>
<td>4</td>
<td>10.977</td>
<td>0.027</td>
<td>29.929</td>
<td>0.000</td>
</tr>
<tr>
<td>12</td>
<td>21.654</td>
<td>0.042</td>
<td>89.690</td>
<td>0.000</td>
</tr>
<tr>
<td>B. Two recurring regimes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.016</td>
<td>0.901</td>
<td>0.049</td>
<td>0.826</td>
</tr>
<tr>
<td>4</td>
<td>11.087</td>
<td>0.026</td>
<td>1.008</td>
<td>0.909</td>
</tr>
<tr>
<td>12</td>
<td>21.369</td>
<td>0.045</td>
<td>19.320</td>
<td>0.081</td>
</tr>
<tr>
<td>C. Two permanent regimes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.092</td>
<td>0.761</td>
<td>0.000</td>
<td>0.996</td>
</tr>
<tr>
<td>4</td>
<td>11.435</td>
<td>0.022</td>
<td>1.102</td>
<td>0.894</td>
</tr>
<tr>
<td>12</td>
<td>21.441</td>
<td>0.044</td>
<td>19.310</td>
<td>0.081</td>
</tr>
<tr>
<td>D. Three recurring regimes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.381</td>
<td>0.527</td>
<td>0.001</td>
<td>0.977</td>
</tr>
<tr>
<td>4</td>
<td>5.447</td>
<td>0.244</td>
<td>1.429</td>
<td>0.839</td>
</tr>
<tr>
<td>12</td>
<td>18.357</td>
<td>0.105</td>
<td>16.487</td>
<td>0.170</td>
</tr>
<tr>
<td>E. Three permanent regimes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.965</td>
<td>0.326</td>
<td>0.456</td>
<td>0.499</td>
</tr>
<tr>
<td>4</td>
<td>6.231</td>
<td>0.183</td>
<td>1.923</td>
<td>0.750</td>
</tr>
<tr>
<td>12</td>
<td>18.113</td>
<td>0.112</td>
<td>18.715</td>
<td>0.096</td>
</tr>
</tbody>
</table>

Note: The \( Q^* \)-statistic refers to the modified Ljung-Box portmanteau test statistic.
very small given the number of restrictions, although it does not have a standard
distribution as the restrictions are on the boundary of the parameter space.
Meanwhile, we verify the model selection results with the residual diagnostic tests
in Table 2. The test results support using a version of the model with three regimes
instead of two regimes, as three regimes are necessary to capture all of the serial
correlation and heteroskedasticity in the data.

4.2. Estimates for the Unobserved Components Model and the Timing of Regime
Changes

Table 3 reports maximum likelihood estimates for the UC model with three
permanent regimes and an AR(2) transitory component. The first results to notice
are for the standard deviations of permanent and transitory shocks in each regime.
In terms of the permanent shocks, the estimated standard deviation is greater than
zero in each regime, but is much larger in the second regime (S_t = 2) than in the
first regime (S_t = 1) or third regime (S_t = 3). Thus, consistent with Stock and
Watson (2007), the U.S. inflation rate appears to have been subject to permanent
shocks over the entire sample period, but the size of permanent shocks has
changed over time. Meanwhile, in terms of the transitory shocks, the estimated
standard deviation has changed very little over the sample period, which is also
consistent with the findings in Stock and Watson (2007) using stochastic volatility.
Thus, the ratio of standard deviations for permanent and transitory shocks, which
is the crucial determinant of the persistence of inflation over long horizons,
appears to have changed due to changes in the size of permanent shocks. The
estimated ratios of standard deviations for permanent and transitory shocks are
0.73, 2.69, and 0.65 in the three respective regimes. Conditional on the

8 In preliminary analysis, we considered a range of specifications for our UC model with different
ARMA processes for the transitory component and different numbers of regimes. We found that,
regardless of the number of regimes, an AR(2) specification was sufficient to capture the dynamics
of the transitory component.

9 Nason (2006) also finds instability in the relative importance of permanent shocks to inflation
using rolling-sample estimates for a UC model of U.S. inflation over the sample period of 1967-
2005. Similarly, Piger and Rasche (forthcoming) find multiple structural breaks in the variance of
a time-varying intercept in a Phillips Curve regression equation for the United States over the
assumption of regime switching, we can use a standard likelihood ratio test to determine whether these ratios are significantly different across regimes. The test statistic is 6.69, which has a $p$-value of 0.04 based on a $\chi^2(2)$ distribution under the null hypothesis that the ratios are the same in each regime (i.e., there is no change in persistence over long horizons). Thus, we are able to reject the hypothesis that persistence has remained constant.

The other results to notice in Table 3 are for the regime continuation probabilities. Corresponding to the restrictions in (8'), the first two regimes are “terminal” and the last regime is “absorbing”. Therefore, there are only two free probabilities to be estimated: $p_{11}$ and $p_{22}$. The estimates for these parameters

<table>
<thead>
<tr>
<th>Table 3 - Maximum Likelihood Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>$S_t = 1$</td>
</tr>
<tr>
<td>$\mu_{t,S}$</td>
</tr>
<tr>
<td>0.0996 (0.0368)</td>
</tr>
<tr>
<td>$\phi_{t,S}$</td>
</tr>
<tr>
<td>0.2015 (0.1513)</td>
</tr>
<tr>
<td>$\sigma_{\eta,S}$</td>
</tr>
<tr>
<td>0.2244 (0.0751)</td>
</tr>
<tr>
<td>$\sigma_{\xi,S}$</td>
</tr>
<tr>
<td>0.5098 (0.0963)</td>
</tr>
<tr>
<td>$\rho_{S}$</td>
</tr>
<tr>
<td>0.0699 (0.9974)</td>
</tr>
<tr>
<td>$p_{11}$</td>
</tr>
<tr>
<td>0.9786 (0.0212)</td>
</tr>
<tr>
<td>Log Likelihood</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.
indicate that each regime was very persistent. Figure 1 confirms this result by displaying the smoothed probability of each regime, $\Pr[S_t = i|\Omega_T]$, computed using Kim’s smoothing algorithm (see Kim and Nelson, 1999).

Given the estimates of the continuation probabilities, $\hat{p}_{11}$ and $\hat{p}_{22}$, we can
estimate the expected duration (in quarters) for each regime as

\[ \hat{D}_1 = \frac{1}{1 - \hat{p}_{11}} \quad \text{and} \quad \hat{D}_2 = \frac{1}{1 - \hat{p}_{22}} \]

Then, the first and the second structural breakpoints can be estimated as \( \hat{D}_1 \) and \( (\hat{D}_1 + \hat{D}_2) \) quarters after the beginning of the sample in 1959Q1, corresponding to 1970Q3 and 1984Q4, respectively, as shown in Table 4. Furthermore, by differencing the smoothed probabilities, we can obtain posterior densities for the structural breakpoints, as plotted in Figure 2. The regimes and their transitions are well identified, with the first regime lasting from the beginning of the sample in 1959 until the early 1970s, the second regime lasting from the early 1970s until the mid-1980s, and the third regime lasting from the mid-1980s through the end of the sample in 2006. It is important to notice that the posterior densities for the breakpoints are fairly sharp, consistent with the idea that the structural changes were sudden rather than gradual. In particular, if the underlying structural changes were more gradual, as would be assumed with a stochastic volatility specification (e.g., Stock and Watson, 2007), we would expect to see less precision for the timing of the breakpoints when using a Markov-switching model.

4.3. Estimates for the Reduced-Form Model and the Persistence of Inflation

We formally investigate changes in inflation persistence using regime-dependent impulse response functions. In order to capture the overall impact of the size and propagation of shocks on the persistence of the inflation process, we transform the UC model of inflation in Table 3 into its corresponding reduced-form

<table>
<thead>
<tr>
<th>Structural Breakpoints and Credibility Bands</th>
<th>Estimated Date</th>
<th>95% Credibility Bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Structural Breakpoint</td>
<td>1973Q3</td>
<td>1968Q4 – 1972Q1</td>
</tr>
<tr>
<td>2nd Structural Breakpoint</td>
<td>1984Q4</td>
<td>1982Q2 – 1986Q4</td>
</tr>
</tbody>
</table>

Notes: Estimates are based on the expected duration of each regime. Credibility bands are based on smoothed probabilities conditioned on estimated parameters.
ARIMA(2,1,2) model. Table 5 reports the implied estimates for the reduced-form model. Using these estimates, we calculate the regime-dependent cumulative impulse response functions for inflation. Figure 3 displays the impulse response functions for a one-unit forecast error. The most notable feature of the results is that, given the same initial impact, the long-run response in the second regime is much larger than in the other regimes. From the 1.96 standard error bands (based on the delta method), we can see that the long-run response is only significant for the second regime. Overall, our estimates confirm that inflation was more persistent over all horizons between the early 1970s and mid-1980s than it was before or after. Meanwhile, inflation appears to be somewhat more persistent after the mid-1980s than it was in the 1960s, explaining the need for three distinct regimes instead of only two regimes with the low persistence regime reoccurring.10

In terms of the gap between inflation and its long-run trend, we can look directly at the autoregressive parameters in Table 3 (or Table 5) to measure persistence. Because the inflation gap is stationary, it makes more sense to look at

Fig. 2
U.S. Inflation and Posterior Densities for Structural Breakpoints

10 It should also be noted that the first and third regimes differ in terms of the dynamics of inflation. In particular, the drift term in the first regime is significantly positive, while it is not significant in the third regime. In other words, there was positive drift in inflation during the 1960s, but not in the late-1980s or afterwards.
the implied half-life response of the inflation gap, rather than the cumulative long-run response. Notably, the half-life response to a shock is only one quarter for each regime.\textsuperscript{11} Thus, it appears that the persistence of the inflation gap remained very low throughout the entire sample period.

5. Conclusion

By estimating an unobserved components model of inflation with Markov-switching parameters and solving for the impulse response function for inflation in each regime, we have found evidence for two permanent regime shifts in U.S. inflation persistence over the period of 1959-2006. Specifically, the breakpoints were estimated to be 1970Q3 and 1984Q4, respectively, with the structural

\textsuperscript{11} We define a half-life as the number of quarters after which the impulse response function for a one-unit shock remains below 0.5.
changes appearing to be very sudden.

The timing of these regime shifts is highly suggestive of a link between monetary policy regimes and the persistence of inflation. In particular, the first regime shift in inflation persistence coincided with the collapse of the Bretton
Woods system in the early 1970s, with persistence rising due to an increase in the variance of permanent shocks to inflation. This timing matches with the notion that U.S. monetary policy was conducted without adherence to the Taylor principle during the 1970s (see, for example, Clarida, Gali, and Gertler, 2000), implying a lack of a stable long-run anchor for the level of inflation. Meanwhile, the second regime shift in inflation persistence coincided with Paul Volcker’s successful taming of inflation expectations by the mid-1980s, with persistence declining due to a decrease in the variance of permanent shocks to inflation. This timing matches with the notion that monetary policy in the Volcker-Greenspan era conformed to the Taylor principle, implying a relatively stable level of inflation in the long run. Importantly, unlike demographic changes that could also have produced the same broad pattern in the level of inflation due to time-consistency problems (see Ireland, 1999), the changes in monetary policy regimes were fairly discrete. Thus, our finding of sudden regime shifts in inflation persistence supports the idea argued for by Benati (2008) that the degree of inflation persistence depends closely on the presence or absence of a well-defined nominal anchor for monetary policy. Meanwhile, the lack of persistence in the gap of inflation from its long-run trend provides some support for the purely forward-looking version New Keynesian Phillips Curve in which prices are sticky, but the inflation gap is not. Of course, the relative importance of changes in trend

12 Under the Taylor principle, the policy interest rate is set to respond more than one-for-one to a movement in inflation away from a long-run target. In a recent paper on this topic, Davig and Doh (2009) estimate a DSGE model with sticky prices and Markov-switching policy regimes. In a two-regime version of their model, they find timing for a “passive” policy regime (i.e., the response to inflation is less than one-for-one) that coincides almost exactly with the timing of our second regime (i.e., the regime with the highest inflation persistence). By linking the response of monetary policy to structural shocks with different persistence, they also find timing for changes in overall inflation persistence for a four-regime version of their model that is broadly similar to what we find, with persistence increasing in the 1970s and decreasing in the early 1980s. Meanwhile, in another recent paper, Murray, Nikolsko-Rzhevskyy, and Papell (2009) estimate Markov-switching models for inflation and the Taylor rule. They find a similar “low-high-low” pattern for inflation persistence to what we have found, although they measure persistence using the sum of the autoregressive coefficients for a Markov-switching AR model of inflation. They also find regime shifts in their estimated Taylor rule that roughly conform to the timing of our regime shifts in inflation persistence, with a passive policy regime prevailing between 1973Q1-1975Q1 and 1979Q4-1985Q3.
inflation over the past half-century means that the New Keynesian Phillips Curve provides a theoretical explanation for a relatively small portion of the overall movements in inflation.

References


