Title	Has Euro-area inflation persistence changed over time?
Authors(s)	O'Reilly, Gerard, Whelan, Karl
Publication date	2005-11
Publication information	O'Reilly, Gerard, and Karl Whelan. "Has Euro-Area Inflation Persistence Changed over Time?" 87, no. 4 (November, 2005).
Publisher	MIT Press
Item record/more information	http://hdl.handle.net/10197/211
Publisher's version (DOI)	10.1162/003465305775098125

HAS EURO-AREA INFLATION PERSISTENCE CHANGED OVER TIME?

Gerard O'Reilly and Karl Whelan*

Abstract—This paper analyzes the stability over time of the econometric process for euro-area inflation since 1970, focusing in particular on the behavior of the so-called persistence parameter (the sum of the coefficients on the lagged dependent variables). Perhaps surprisingly, in light of the Lucas critique, our principal finding is that there appears to be relatively little instability in the parameters of the euro-area inflation process. Full-sample estimates of the persistence parameter are generally close to 1, and we fail to reject the hypothesis that this parameter has been stable over time. We discuss how these results provide some indirect evidence against rational expectations models with strong forward-looking elements, such as the New Keynesian Phillips curve.

I. Introduction

The European Central Bank has an explicit mandate for the maintenance of low inflation as its overriding objective. In light of this mandate, an obvious goal for macroeconomists wishing to analyze European monetary policy is the development of statistically adequate econometric models of the euro-area inflation process. However, though such a goal is clear in principle, the task may be complicated by a number of practical problems. Firstly, there are the potential problems due to modeling a series that aggregates the inflation processes of various countries that have historically pursued independent monetary policies. In addition, the substantial changes over time in monetary regimes may leave any econometric model of euro-area inflation open to the Lucas critique. In other words, given the sequence of shifts in monetary policy regimes that have occurred since the early 1970s, it would hardly be surprising if euro-area inflation regressions exhibited substantial parameter instability, rendering them of dubious usefulness for forecasting or policy analysis. A particular concern about these regressions that has emerged in recent years, as researchers have increasingly used forward-looking rational-expectations models of inflation such as the New Keynesian Phillips curve, is the idea that the importance of lagged dependent-variable terms should decline as the credibility of a central bank's commitment to low inflation increases; this theme has been emphasized by John Taylor (1998), Thomas Sargent (1999), and others.

With this background in mind, this paper analyzes the stability over time of some simple econometric representations of the euro-area inflation process since 1970, focusing in particular on the behavior of the so-called *persistence parameter,* which is defined as the sum of the coefficients on the lagged dependent variables. In this respect, our paper adds to a recent literature that has been devoted to documenting the facts in relation to structural changes over time in the persistence of various inflation processes, including Cogley and Sargent's (2002), Stock's (2002), and Pivetta and Reis's (2003) studies of this issue for the United States, and Levin and Piger's (2003) multicountry study. More generally, because the euro-area inflation process provides a clear example of a region and a series to which the Lucas critique is most likely to apply, we believe our analysis provides some useful evidence for assessing the empirical importance of changes in policy regimes for the parameters of reduced-form macroeconometric processes.

Perhaps surprisingly, given the theoretical priors just outlined, our principal finding is that there appears to be relatively little instability in the parameters of the euro-area inflation process. Our full-sample estimates of the persistence parameter are generally close to 1, and results from unknown-breakpoint tests for structural change are consistent with the null of no change over time in this coefficient. These tests do appear to detect a structural break in the intercept term, and conditioning on such a break produces somewhat lower estimates of the persistence parameter. However, we show below that the standard asymptotic p-values used to implement these unknown-breakpoint tests turn out to be poor approximations to the correct finite-sample distributions when the true value of the persistence parameter is close to or equal to 1. We also show that once this factor is corrected for, there is no significant evidence of an intercept break.

Of course, the failure to formally reject a null hypothesis of parameter stability does not, on its own, rule out the existence of some important structural changes. In light of this possibility, we also report results from rolling regressions, which allow for separate parameters for the inflation process over a sequence of moving windows. These exercises show that estimates of the persistence parameter are relatively stable throughout the estimation period, with our preferred point estimates usually being very close to 1. These preferred estimates were obtained using Bruce Hansen's (1999) grid-bootstrap procedure, which corrects for the finite-sample biases that occur with OLS estimation. It is also worth noting that although most of our exercises follow the existing literature in focusing on univariate regressions, our conclusions concerning the persistence parameter are also robust to specifications that include an output gap. Overall, our results are consistent with a stable reduced-form representation for inflation and a high level of inflation persistence.

This finding of a high and stable persistence parameter may be somewhat surprising, given the obvious potential relevance of the Lucas critique for our exercise. However, it

Received for publication March 11, 2004. Revision accepted for publication November 21, 2004.

* Central Bank and Financial Services Authority of Ireland.

This paper reflects research conducted as part of the Eurosystem Inflation Persistence Network. We wish to thank participants in the network, as well as Jim Stock and two referees, for helpful comments. We also wish to thank Jerome Henry for generously providing us with the data used in this study. The views expressed in this paper are our own, and do not necessarily reflect the views of the Central Bank and Financial Services Authority of Ireland or the ESCB.

The Review of Economics and Statistics, November 2005, 87(4): 709-720

is consistent with recent evidence for the United States presented by Rudebusch (2005), who uses an estimated New-Keynesian-style macroeconomic model to show that the parameters of reduced-form regressions will tend to be relatively stable even in the presence of realistic changes in monetary policy rules. Rudebusch also shows that if the underlying structural equations of such models place relatively low weights on forward-looking expectational variables, then the inflation persistence parameter in reduced-form models will be close to 1. Thus, we interpret our results as providing some indirect evidence against pricing models with strong forward-looking elements, such as the New Keynesian Phillips curve.

II. Theoretical Background and Policy Implications

Since the seminal works of Friedman (1968) and Lucas (1972a), it has become widely accepted that the behavior of the aggregate inflation process depends crucially on those factors that influence the expectations of private agents. As a result, almost all brands of theorizing about inflation now emphasize the important role played by expectations. For example, textbook treatments of inflation such as Blanchard (2000) focus on the role played in wage and price setting by workers' prior expectations of price inflation, implying a specification of the form

$$\pi_t = E_{t-1}\pi_t + \eta y_t + \epsilon_t, \tag{1}$$

where π_t and y_t represent the inflation rate and output gap, respectively. The modern New Keynesian Phillips curve also emphasizes the importance of expectations, although the mechanisms through which this operates are somewhat different. For example, Calvo-style models feature rational price setters whose concern about their future margins requires them to consider future inflation when setting prices that may be fixed for a number of periods. This results in an inflation equation of the form

$$\pi_t = \beta E_t \pi_{t+1} + \gamma y_t, \tag{2}$$

where β is a discount rate close to 1.[1]

Crucial differences emerge, however, once one turns to the empirical modeling of these expectations. Reduced-form Phillips curves are commonly estimated as

$$\pi_t = \alpha + \rho\pi_{t-1} + \sum_{k=1}^{n} \psi_k \Delta\pi_{t-k} + \delta Z_t + \epsilon_t, \tag{3}$$

where Z_t is a vector of other variables that may affect inflation. The motivation for this specification is that agents extrapolate from past inflation rates to formulate the expectation used in current-period wage and price setting. Frequently, the value of ρ is restricted to equal 1 implying that agents formulate rule-of-thumb expectations based on a weighted average of past inflation rates; this specification is often motivated by a desire to rule out a long-run tradeoff between the levels of inflation and output.

If econometric equations such as (3) are relatively stable over time, then the lagged dependent-variable terms are of great importance for the design of monetary policy. In this case, these terms describe how shocks to inflation today—including those that originate from policy actions—are propagated over time. These considerations suggest that it is crucial that central banks take the estimated persistence parameter into account when setting policy. However, despite their continuing empirical popularity, the theoretical underpinnings of reduced-form Phillips-curve regressions have been in question ever since the early 1970s saw the advent of the rational-expectations approach to macroeconomics. Advocates of this approach emphasized that the type of weighted-average "adaptive" expectation formation underlying these specifications was not necessarily consistent with optimal behavior.[2]

For our analysis of the euro-area inflation process, a particular concern posed by the assumption of rule-of-thumb expectations is that this type of model may work poorly in a world in which central-bank behavior changes over time, as stressed in Lucas's (1976) famous critique of econometric modeling.[3] The extrapolation of expected inflation based only on past values may be reasonable if a central bank allows its target rate of inflation to drift over time, but a switch to a credible low-inflation target may make such a rule less sensible. For example, if it is known that a central bank has a credible commitment to a 2% target for inflation each period, then it may be rational for agents to always expect inflation to be around 2%, implying a reduced-form inflation process approximately of the form

$$\pi_t = 2 + \delta Z_t + \epsilon_t. \tag{4}$$

This type of process rules out a role for the lagged inflation terms altogether.

With the rational-expectations-based New Keynesian Phillips curve playing an increased role in monetary policy analysis in recent years, this particular idea—that a credible commitment to low inflation will lead to a reduction in the estimated value of ρ—has been quite widely discussed of late.[4] This conjecture, of course, has substantial implications for monetary policy: If reduced-form estimates of ρ are indeed

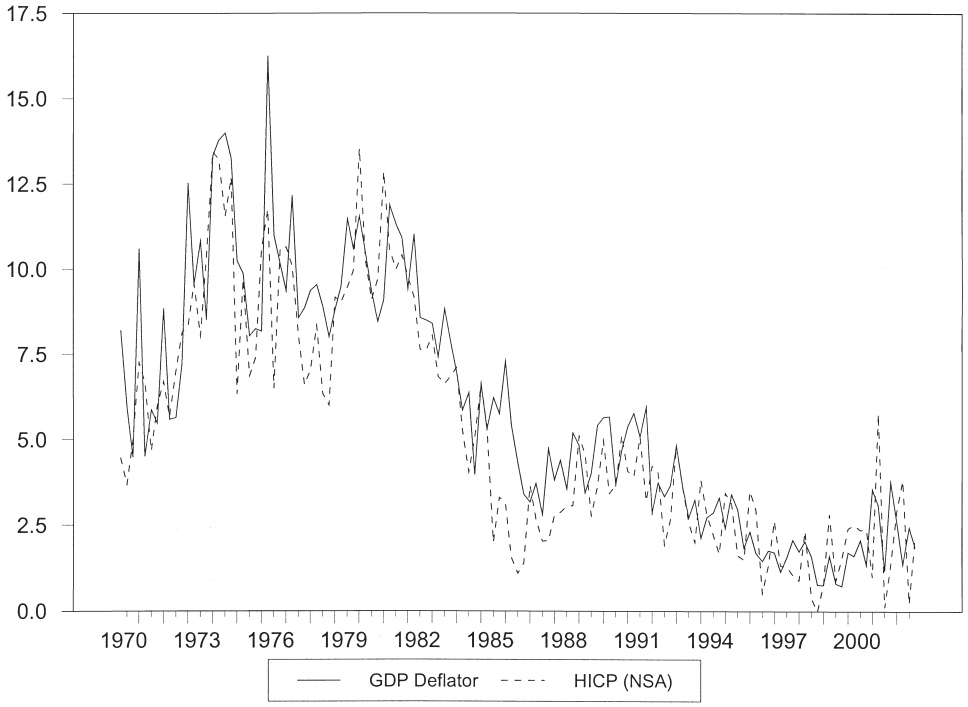
[1] The New Keynesian Phillips curve can also be derived from other microfoundations, such as models featuring costly price adjustment or staggered wages. See Roberts (1995).

[2] For example, Sargent (1971) argued that the U.S. inflation process appeared to be stationary and so the econometric specifications with $\rho = 1$ were at odds with an expectations formation process based on available data at the time. Lucas (1972b) also questioned whether tests of $\rho = 1$ in this equation could correctly be interpreted as tests for monetary neutrality.

[3] Indeed, the effect of changes in monetary policy on inflation expectations was a specific example discussed in Lucas's paper.

[4] See Taylor (1998), Sargent (1999), Cogley and Sargent (2001), and Levin and Piger (2003) for discussions of the likely effect that a credible commitment to a low inflation target has on the parameter ρ. See Clarida, Galí, and Gertler (1999) for an example of monetary policy analysis from the perspective of the New Keynesian Phillips curve.

FIGURE 1.—EURO-AREA PRICE INFLATION (QUARTER-OVER-QUARTER, ANNUAL RATE)



spuriously high or unstable over time, then many of the results from econometric exercises based on such regressions could be spurious, and policy must rely on other, more structural models that are capable of explaining the shifting reduced-form dynamics. And the euro area provides a particularly relevant testing ground for these ideas, given the series of regime changes seen since 1970: The breakup of the Bretton Woods framework, the formulation and gradual hardening of the EMS system, and the run-up to and introduction of the EMU with its strict low-inflation mandate.

III. Full-Sample Results

The data source for our analysis is an updated version of the ECB's Area Wide Model (AWM) quarterly database described in Fagan, Henry, and Mestre (2001). The sample for this data set is 1970:1–2002:4. Our principal inflation measure is the annualized quarterly log-difference of the GDP deflator. We also report some results for the Harmonized Index of Consumer Prices (HICP), the annual change in which is cited in the ECB's official inflation mandate. Importantly, though the AWM series for the GDP deflator is seasonally adjusted, the HICP series is not. These two inflation series are plotted on figure 1.

For our first set of calculations, we follow the approach taken in some other recent studies such as Pivetta and Reis (2003) and Levin and Piger (2003) in focusing on the estimation of the parameter ρ in univariate regressions of the form

$$\pi_t = \alpha + \rho\pi_{t-1} + \sum_{k=1}^{n} \psi_k \Delta\pi_{t-k} + \epsilon_t. \tag{5}$$

We set $n = 3$ (consistent with four lags of the level of inflation) on the basis of lag selection tests, but none of the substantive results that we report were sensitive to this choice. Moreover, because the HICP series is not seasonally adjusted, the inclusion of four lags in the levels specification is appropriate, as this allows us to capture average seasonal patterns in this series.

There are a number of good reasons for focusing on the parameter ρ as our principal measure of inflation persistence. For example, for this model, ρ is a crucial determinant of the response to shocks over time: It can be shown that the infinite-horizon cumulative impulse response (CIR) to shocks (where this is defined) is given by $1/1-\rho$. However, an advantage of focusing on the estimate of ρ rather than on CIRs is that this measure of persistence remains defined even when the underlying process contains a unit root or is explosive. Given that we do not wish to place any prior restrictions on the nature of the inflation dynamics, this is a distinct advantage for the sum-of-the-autoregressive-coefficients measure. Moreover, we can note ahead of time that our conclusions about the behavior over time of euro-area inflation persistence are not determined by this choice of persistence measure, similar conclusions being reached when we use other measures such as the largest autoregressive root or finite-horizon CIRs.[5]

Table 1 reports the OLS estimates of ρ for both deflators. For both series the point estimates were approximately 0.96,

[5] Andrews and Chen (1994) discuss this issue and compare the sum of the coefficients on lagged dependent variables favorably with other popular measures of persistence such as the largest autoregressive root and the half-life.

TABLE 1.—FULL-SAMPLE UNIVARIATE OLS ESTIMATES OF ρ

GDP deflator	0.960 (0.039)
HICP	0.957 (0.043)

Standard errors in parentheses below the coefficients.

TABLE 2.—UNIVARIATE GRID-BOOTSTRAP ESTIMATES

	Median Unbiased	90% Confidence Interval
GDP deflator	1.022	(0.938, 1.060)
HICP	1.025	(0.937, 1.065)

Note: Estimated by Bruce Hansen's (1999) grid-*t* bootstrap method using GAUSS code downloaded from Hansen's Web site.

and the augmented Dickey-Fuller (ADF) statistics associated with these regressions do not come close to rejecting the null hypotheses that the inflation series have unit roots without drift. One problem with these estimates is the fact that, in finite samples, the standard asymptotic distributions for OLS coefficients and *t*-statistics in autoregressive models are known to become systematically poorer approximations to their true finite-sample distributions as the true value of ρ increases. In particular, point estimates become increasingly downward biased and their distribution becomes more skewed to the left as ρ increases.

To rectify this problem, we also used Bruce Hansen's (1999) grid-bootstrap method to obtain bias-adjusted point estimates and confidence intervals.[6] This method uses a bootstrap technique to simulate the finite-sample distribution of the OLS estimator for a range of possible true values of the parameter ρ in the model:

$$Y_t = \delta_0 + \delta_1 t + Z_t, \tag{6}$$

$$Z_t = \rho Z_{t-1} + \epsilon_t. \tag{7}$$

This approach produces median-unbiased estimates of ρ; in other words, it tells us the value of ρ that would result in the estimated OLS parameter $\hat{\rho}$ being the median of the empirical sampling distribution. This method also allows for the construction of confidence intervals that accurately capture the skewed nature of the finite-sample distributions: The upper limit of the grid-bootstrap confidence interval is the value of ρ for which the OLS estimate is the 5th percentile of the sampling distribution, and the lower limit corresponds to the value consistent with the OLS estimate being the 95th percentile of the sampling distribution.

In our implementation, we set the parameter δ_1 equal to 0, as the hypotheses that both inflation series have a unit root with drift can be firmly rejected.[7] Table 2 reports the results from our grid-bootstrap estimation. They show that our OLS estimates are actually consistent with point estimates of ρ of approximately 1.02, with the lower end of the 90% confidence intervals equaling approximately 0.94 for both series. These results enforce the picture painted by the OLS estimates of a highly persistent series: The median-unbiased representation of the euro-area inflation process is essentially one with a unit root without drift, and even the lower ranges of our estimates of ρ are consistent with a high degree of persistence.

IV. Tests for Structural Change

Our full-sample univariate estimates suggest a high level of inflation persistence. However, a number of explanations are possible as to why these high estimates of ρ may be misleading, or possibly completely spurious. The first potential problem, in light of the Lucas critique, is that the assumption of a constant ρ throughout the sample may be inappropriate: Our high full-sample estimate could still mask a substantial reduction in persistence over the latter part of the sample.

Another potential problem is the fact that these calculations do not allow for the possibility of a shift in the unconditional mean value of inflation. The fact that inflation was, on average, high in the early part of the sample and low in the later part implies that allowing for such a shift would improve the fit of the model.[8] Also, it is well known from work such as Perron (1989) that failure to allow for structural breaks in an intercept or trend can result in spuriously high estimates of the persistence parameter: Once one allows for changes over time in the mean, then deviations from this time-varying mean do not seem as persistent. On the other hand, the very fact that inflation was high in one part of the sample and low in another is not, on its own, evidence against a model with a constant unconditional mean. In particular, a constant-mean process with a high value of ρ is quite capable of generating periods of high inflation followed by periods of low inflation. Ultimately, we need to test formally whether the null hypothesis of parameter stability can be rejected.

A. *Tests Based on Asymptotic Distributions*

We are interested in testing the general null hypothesis of parameter stability. Thus, instead of carrying out a traditional Chow test, which posits a specific break date, our structural-change tests do not assume any prior knowledge about potential break dates. Two test statistics were calculated. The first is the Andrews-Quandt sup-*F* statistic, which is the maximum of a sequence of traditional Chow-style χ^2 tests for structural change, each based on a different potential breakpoint. This test statistic was originally introduced by Quandt (1960), and its asymptotic distribution was de-

[6] Hansen's paper actually develops two different grid bootstrap estimators that produce very similar answers. We use his preferred method, which he labels the grid-*t* method. GAUSS code to produce these estimates was downloaded from Hansen's Web site.

[7] In this case, the distribution of the OLS estimate collapses quickly on 1, and our point estimates are far too low to be consistent with this hypothesis. Note that this implies that an estimate of ρ equal to 1 from the grid bootstrap procedure is not necessarily the same as an estimate of 1 from equation (5).

[8] This unconditional mean is estimated as $\alpha/1-\rho$ in our OLS regression, and as the sample-mean inflation rate in the grid-bootstrap estimates.

TABLE 3.—UNKNOWN-BREAKPOINT TESTS FOR STRUCTURAL CHANGE

	ρ-Break	Intercept-Break
Sup-*F* (Andrews-Quandt) Tests		
Test values	6.82	10.47
Asymptotic *p*-values	0.11	0.02
Bootstrapped *p*-values (ρ = 0.96)	0.34	0.20
Bootstrapped *p*-values (ρ = 1.00)	0.37	0.24
Exp-*F* (Andrews-Ploberger) Tests		
Test values	1.38	2.58
Asymptotic *p*-values	0.12	0.03
Bootstrapped *p*-values (ρ = 0.96)	0.36	0.21
Bootstrapped *p*-values (ρ = 1.00)	0.39	0.24

Note: Bootstrapped *p*-values are based on simulating the estimated OLS process (the ρ = 0.96 case) or the estimated process with ρ = 1 imposed; shocks for the simulated processes were based on drawing from the estimated residuals.

rived by Andrews (1993).[9] The second test is the exp-*F* statistic, which is based on a weighted average of the full sequence of χ^2 tests; this test and its asymptotic distribution were introduced by Andrews and Ploberger (1994). Though less commonly used, the exp-*F* has been shown to have superior power in distinguishing the null hypothesis from local alternatives.

In what follows, we only report results for the GDP deflator, rather than testing for parameter stability with the non-seasonally-adjusted HICP data, which may exhibit instabilities over time due to changing seasonal patterns. The test statistics and their asymptotic *p*-values are reported in table 3. In addition, figure 2 plots the time series of Chow statistics associated with the various potential break dates; the left-hand panel illustrates the results for tests of stability of the persistence parameter, and the right-hand panel shows results for the intercept. The figure also includes the relevant 5% critical value for the Andrews asymptotic distribution for the sup-*F* statistic.[10]

Figure 2 shows that the maximum Chow statistic for a break in the persistence parameter is 6.82 (this occurs at 1982:3). Though this break is technically significant for the traditional χ^2 distribution, this value falls well short of the 5% critical value for the Andrews distribution: Using the approximate asymptotic distributions calculated by Hansen (1997), this result has a *p*-value of 11%. (Results for the exp-*F* statistic are similar for each case reported here.)

The right-hand panel of figure 2 shows that there is stronger evidence for a break in the intercept term. The sup-*F* statistic (also reached at 1982:3) is 10.47, which is significant at the 2% level of the Andrews distribution. One unsurprising pattern in light of earlier discussions is that conditioning on the break in the intercept reduces the estimated persistence parameter. Allowing for this break, the OLS estimate drops to ρ = 0.80, the median-unbiased estimate is ρ = 0.86, and the 90% confidence interval is (0.75, 1.00). Importantly, however, allowing for a break in the intercept has no effect on our conclusions regarding the *stability* of the persistence parameter: The asymptotic *p*-value for such a test is 0.19.

B. *Problems with the Asymptotic Distributions*

These results show that one cannot formally reject the hypothesis of no break over time in the persistence parameter, but the 11% *p*-value at least suggests the possibility that there may be such a break. And the results also point to the possibility that the correct estimate of ρ is a good deal lower than our full-sample estimates, once one allows for a structural break in the intercept.[11] One concern, however, about these test results is their reliance on asymptotic distributions that may not be appropriate in finite samples for the autoregressive processes that we are considering. The results of Diebold and Chen (1996) indicate that these concerns are likely to be important in this case. Their research shows that the asymptotic *p*-values for sup-*F* tests for simultaneous breaks in the intercept and persistence parameters of an AR(1) model become less accurate as ρ increases, with tests based on these *p*-values too often rejecting the hypothesis of no structural change.

Diebold and Chen's paper does not discuss finite-sample distributions for separate tests for breaks in the intercept or the persistence parameter. To illustrate the finite-sample properties of these tests, figure 3 reports the results from a Monte Carlo analysis of the true sizes of both the intercept and ρ-break tests using the same number of observations as in our estimation sample (T = 127). We simulated a sequence of univariate processes with increasing values of ρ (ranging from 0.4 to 0.999), and with each process having a unit mean and random shocks drawn from a N (0,0.5) distribution; these values were roughly calibrated to match our estimated process for GDP price inflation, for which the standard deviation of the errors was approximately half the model's implied long-run average inflation rate.[12] For each value of ρ considered, we constructed 5,000 separate simulated autoregressive series, and for each of these series we performed sup-*F* tests of the (correct) null hypotheses of parameter stability for the intercept and for the persistence parameter.

Figure 3 shows how, for these parameter values, the empirical size of the sup-*F* test procedures with nominal size of 10% is close to 10% until the true ρ reaches

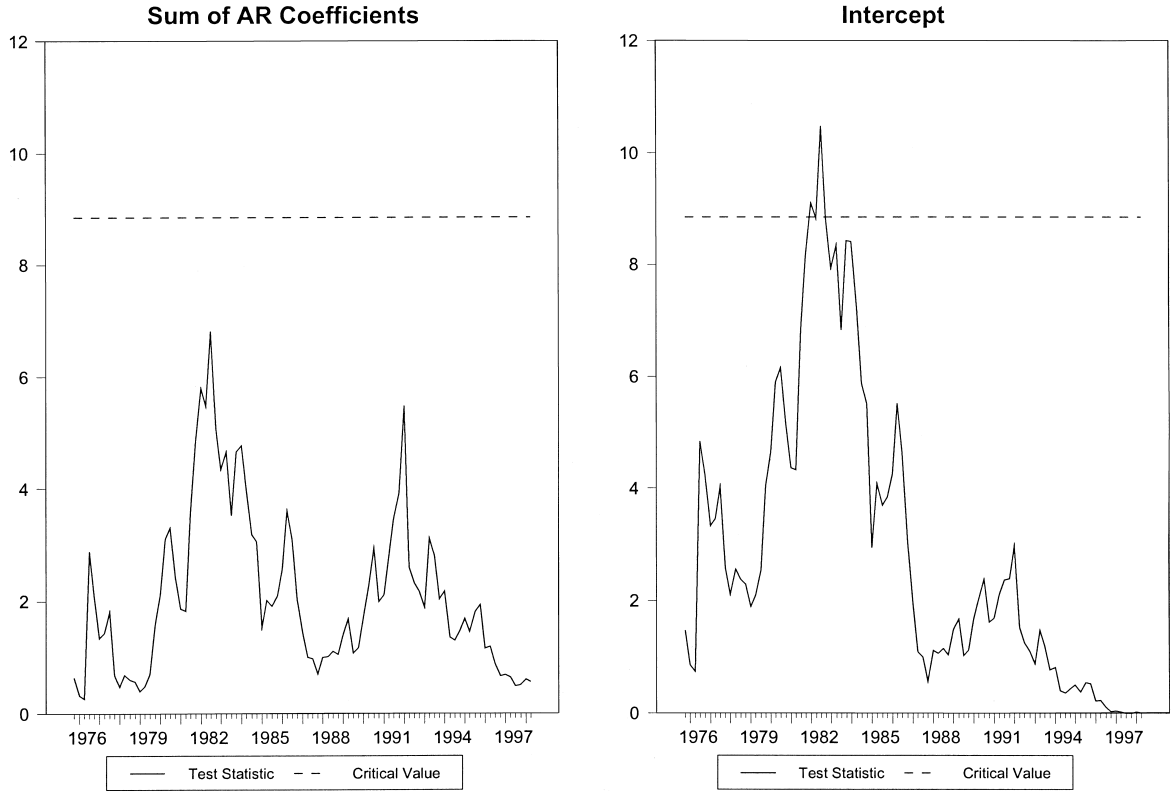
[9] Our analysis in this section uses the Lagrange multiplier (LM) form of the χ^2 test, but similar conclusions are reached using the alternative Wald or likelihood ratio tests.

[10] We followed the usual convention and eliminated the first and last 15% of the observations from consideration as potential breakpoints.

[11] One way to measure the change in persistence caused by allowing for an intercept break is to note that the half-lives associated with the OLS estimates of 0.96 and 0.80 are 17 and 3 quarters, respectively. That said, there are drawbacks to the applicability of the half-life measure in this case. This is because the median-unbiased estimates for these two cases are 1.02 and 0.86, so the half-life is not defined for the best available estimate of the no-break value of ρ. Also, a value of ρ = 1 is still inside the 90% confidence interval even when a break is allowed for.

[12] In other words, we simulated equations (6) and (7) with $\delta_0 = 1$, $\delta_1 = 0$, and $\epsilon_t \sim N(0, 0.5)$.

FIGURE 2.—CHOW TEST SEQUENCES AND 5% CRITICAL VALUES: UNIVARIATE MODEL



approximately 0.7. However, after ρ gets larger than 0.7 the empirical size of the tests increases. The size of the intercept-break test rises particularly rapidly, reaching values of over 50% when ρ is close to 1. In contrast, the size of the ρ-break test rises more slowly until ρ is approximately 0.95, after which it increases rapidly to just under 50% for values of ρ that are very close to 1. It is also noteworthy that these results relate to the LM version of the sup-F test, and that the size distortions will be even greater than those reported here for the alternative Wald version of the test, which is often used in practice.

Given that our full-sample point estimates for ρ are so high, these results imply that using asymptotic distributions overstates the evidence for structural breaks in the euro-area inflation process, perhaps by a significant amount. As an alternative, Diebold and Chen suggested calculating bootstrapped p-values, based on simulating the estimated full-sample process with shock terms drawn randomly from the historical residuals. They show that this technique produces test procedures that have approximately the correct size. We first applied this technique using the OLS point estimates of the GDP inflation process and performing 5,000 bootstrap replications; the results are reported on the third and seventh lines of table 3. The bootstrapped p-value for a sup-F test statistic of 6.82 (the value for a break in ρ) is 34%, not the 11% implied by the asymptotic distribution. Similarly, the bootstrapped p-value for a test statistic of 10.47 (the value for a break in the intercept) is 20% not the 2% reported above. Similar results are obtained for the exp-F statistic.

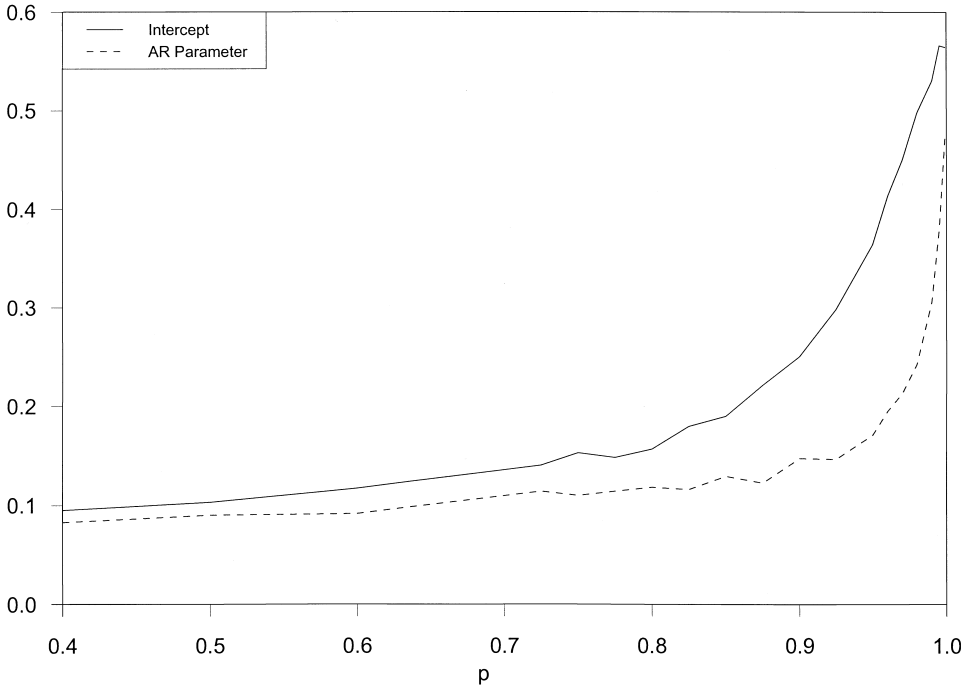
Finally, as discussed above, once one adjusts for the finite-sample bias in the OLS estimates, the true inflation process is well described as a unit root without drift. For this reason, we also performed bootstrap calculations based on simulating the process obtained by estimating the inflation regression with the restriction $\rho = 1$ imposed. This resulted in p-values of 37% for the break in the persistence parameter and 24% for the break in the intercept.

We conclude from these calculations that the evidence from formal hypothesis tests for structural breaks in the euro-area inflation process is quite weak, and the hypothesis that the process has a stable representation with a high level of persistence is hard to reject. In addition, even if one accepts the potential break in the intercept, the value of ρ obtained after conditioning on this break is still fairly high—recall that the median-unbiased estimate of the persistence parameter in this case was 0.86.

C. *Power Considerations*

An important possible objection to the tests just presented is that they take the hypothesis of parameter stability as their null. But even if one fails to formally reject this null, that doesn't necessarily imply that such change is not present. From our perspective of assessing whether or not there have

FIGURE 3.—EXAMPLES OF EMPIRICAL SIZES OF SUP-F 10% TESTS (5,000 DRAWS, $T = 127$)



been important changes over time in inflation persistence, this issue can be characterized in terms of the possibility that one could obtain sup-F statistics as low as we have found even when there is a break in persistence.

Unsurprisingly, calculations designed to assess the power of our test procedures indicate that they have little power to detect small breaks in ρ. More importantly, however, they suggest our tests results are inconsistent with even moderate changes in persistence over time. For example, calculations based on 5,000 replications show that if there is a break halfway through the sample from our full-sample OLS estimate of $\rho = 0.96$ to $\rho = 0.7$, then the probability of obtaining a sup-F test statistic for a break in the persistence parameter as low as our 6.82 is only 12%. And this probability falls to 1% in the case where the break is from $\rho = 1.0$ to $\rho = 0.7$. If we consider a break in which $\rho = 0.6$ over the second half of the sample, then these two probabilities fall to 3% and 0.33%, respectively.

V. Rolling Regressions

In this section, we provide another, more informal method of assessing whether or not there has been structural change over time in the persistence parameter. Specifically, we follow the approach of Pivetta and Reis (2003) and report estimates from a sequence of short rolling samples. Though the small sample sizes involved in these rolling regressions usually imply substantial variation in parameter estimates and wide confidence intervals, they have the advantage of allowing for greater flexibility in detecting structural changes over time, with each rolling sample allowed to have a completely different estimated inflation process.

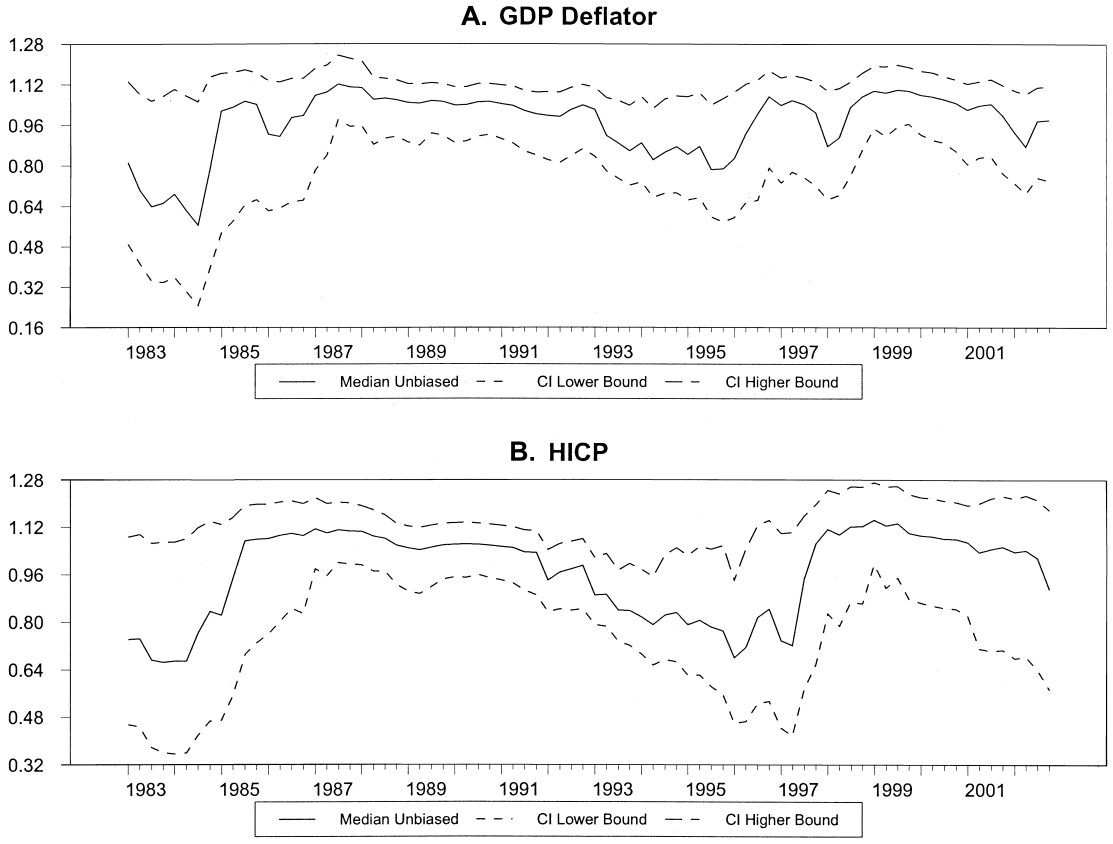
This feature of rolling regressions is particularly likely to be an advantage if one views the high full-sample estimates of ρ as due to a failure to capture time variation in the conditional mean of the inflation process of a more sophisticated type than the once-off breaks considered by the Andrews-Quandt test.[13] For example, although our full-sample estimate of ρ is high, indicating little tendency of inflation to revert to its full-sample mean, it may be that once one considers a sequence of small samples—each more likely to be associated with a specific stable policy regime—then a form of conditional mean reversion could become more apparent.

Figure 4 reports results from these rolling regression exercises. In each case, we estimated an AR(4) model for a sequence of rolling samples and calculated the median-unbiased estimates of ρ and a 90% confidence interval using the grid-bootstrap procedure. Because of the small sample sizes involved in these rolling regression calculations, we believe it is particularly important to focus on median-unbiased point estimates rather than the OLS estimates. This is because it is well known that the finite-sample biases for OLS estimates of autoregressive models get larger as the sample size declines: Whereas tables 1 and 2 report full-sample estimates of the OLS bias for this data set to be approximately 0.06, the bias estimates for the rolling regressions were generally much larger and were often in the range of 0.20 to 0.30.

Despite the intuition discussed above, the rolling regressions generally endorse our earlier conclusion of a high and

[13] See, for example, the discussion of this position by Marques (2003).

FIGURE 4.—ROLLING 12-YEAR GRID-BOOTSTRAP ESTIMATES OF ρ (WITH 90% CONFIDENCE INTERVALS)

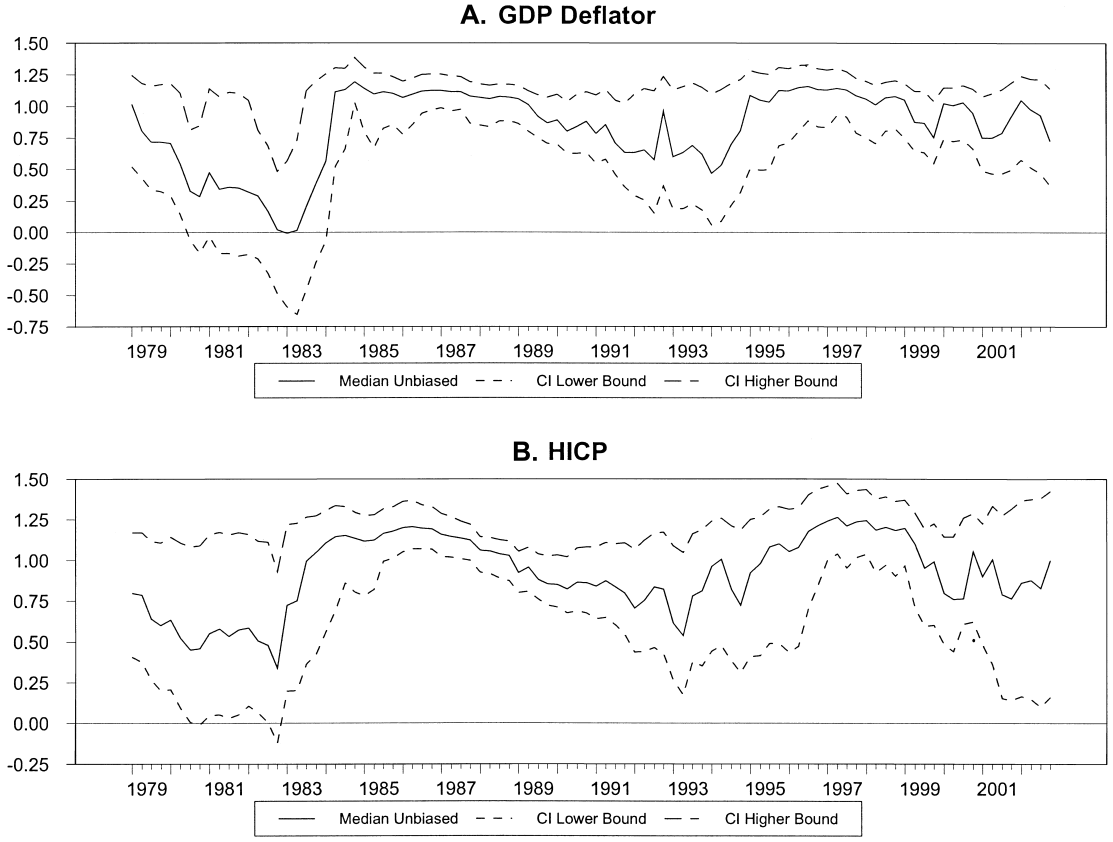


stable value for the persistence parameter. Figure 4A shows the results for the GDP deflator with a rolling window of 12 years. The median-unbiased estimates of the persistence parameter tend to be consistently high: The average value for these estimates is 1.02, and three-quarters of the estimates are greater than 0.89. The upper end of the 90% confidence intervals are stable at a high level, and are always above 1. And despite a dip in the median-unbiased estimates corresponding to samples ending in the early to mid-1990s, there is little evidence of a trend toward lower levels of persistence over time. Figure 4B shows that roughly the same patterns emerge when we use the HICP.

A potential criticism of these results is that the 12-year sample is still too long to capture the behavior of inflation over a single stable monetary policy regime, given the frequency of changes in monetary policy regimes seen in Europe over this period. To address this issue, figure 5 reports results using an 8-year window. As would be expected given the very short estimation windows being used, these results exhibit more volatility than the 12-year results and have wider confidence bands. However, the same overall story emerges: Median-unbiased estimates of ρ are still high in most cases, with no evidence of a tendency toward lower estimates toward the end of the sample. In fact, if anything, these results suggest some increase in persistence after the early 1980s. Consider figure 5A for GDP price inflation: After a sequence of low values associated with samples ending up to 1983, the average value for the median-unbiased estimate for subsequent samples is 1.02.

One methodological point worth noting about these calculations is that they help to illustrate the value of reporting a sequence of estimates from rolling samples, rather than drawing conclusions based on short individual samples. The variability of the estimates from the 8-year samples is very high, and a number of the individual estimates would be quite misleading if presented in isolation. Thus, for example, Kieler (2003) reports an OLS estimate of ρ of 0.55 for the period 1995–2002 from an AR(4) regression for the euro-area GDP deflator, and contrasts this with a full-sample estimate of 0.96 to argue that there has been a sharp decline in inflation persistence in recent years. With our data (which are slightly more up to date), we obtain a very similar OLS estimate for this sample: 0.59. However, as can be seen from the last data points in figure 5(a), this sample produces a substantially higher median-unbiased estimate of 0.72, and the upper end of the 90% confidence interval for this sample is 1.14. And as figure 5A also clearly illustrates, this final rolling sample produces an estimated persistence parameter that is lower than most of the samples close to it, and there is actually little trend over this period toward systematically lower estimates of ρ.

FIGURE 5.—ROLLING 8-YEAR GRID-BOOTSTRAP ESTIMATES OF ρ (WITH 90% CONFIDENCE INTERVALS)



Finally, figure 6 shows that our conclusions about the lack of changes over time in inflation persistence do not depend on our use of the sum of the autoregressive coefficients as our measure of persistence. The figure shows that when persistence is measured based on a median-unbiased estimate of the largest autoregressive root, then the pattern of high and relatively unchanged persistence over time is still apparent.[14] We also found a similar pattern when we measured persistence based on estimated cumulative impulse responses to shocks at medium-term horizons: For instance, our time series of rolling 12-year median-unbiased estimates of the sum of the autoregressive coefficients has a correlation of 0.93 with a time series of rolling estimates of the cumulative impulse response at a 5-year horizon.[15]

VI. Including the Output Gap

Up to this point, we have followed a number of other recent studies in focusing on measuring the persistence parameter in the univariate inflation process. However, as we noted above, practical implementations of econometric Phillips curves usually include some proxy for the level of "slack" in the economy, such as an output gap. And, from our perspective of assessing the level of inflation persistence, there are a number of reasons why we might wish to include such a variable. One simple reason is suggested by the model described by equation (1) augmented with the traditional assumption that expected inflation is a weighted average of past realized values. To the extent that there is negative feedback from inflation to the output gap—for example, because the central bank operates according to an inflation-targeting Taylor rule—univariate exercises will underestimate the true "structural" value of ρ suggested by this model.

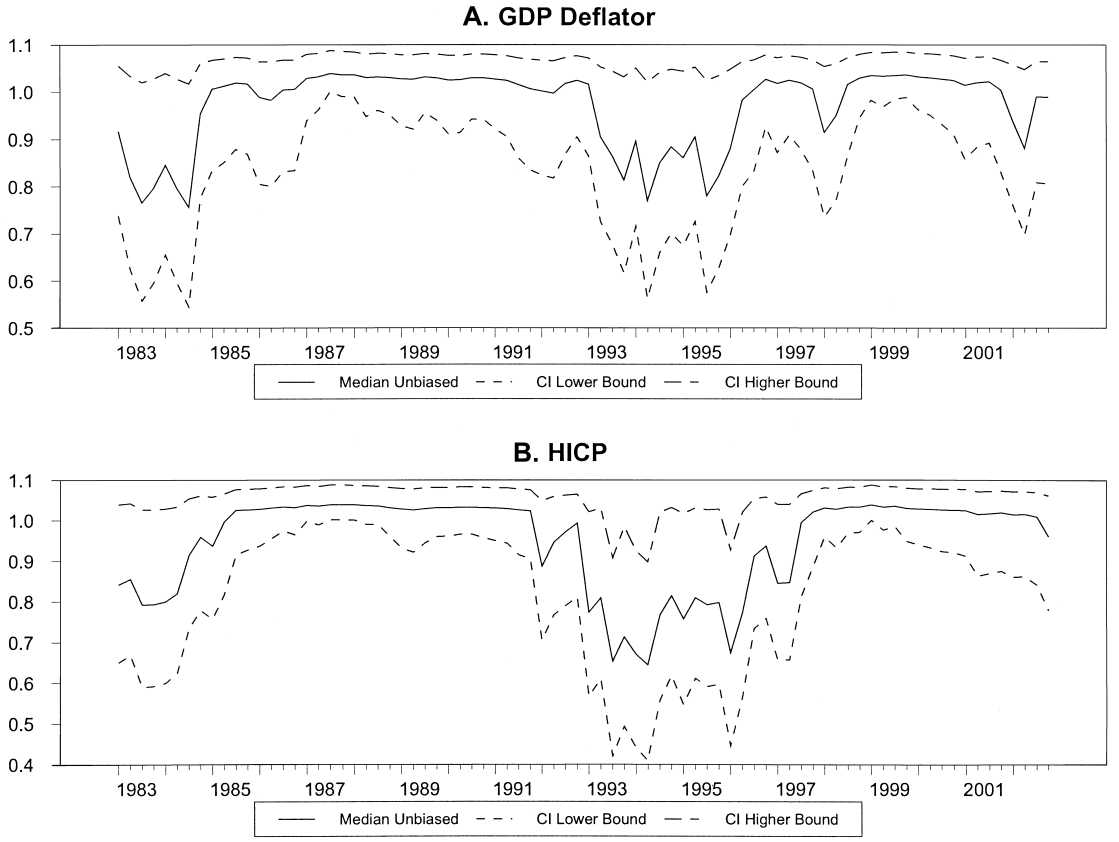
Conversely, it is also possible that the exclusion of an autocorrelated driving variable could result in spurious findings of a high value of ρ. Although evidence of a value of ρ close to 1 might be considered evidence in favor of the "adaptive expectations" approach, models based purely on rational expectations can also predict high values of the persistence parameter in univariate regressions. For example, consider the case in which an output gap follows an AR(1) process

$$y_t = \phi y_{t-1} + u_t. \tag{8}$$

[14] The median-unbiased estimates and 90% confidence intervals in this chart were calculated using the tables in Stock (1991).

[15] One complication, however, with the cumulative-impulse-response measure of persistence is that it makes little sense if the sum of the autoregressive coefficients is greater than 1. Thus, we based these calculations on OLS regression estimates for which the sum of the coefficients was always less than 1.

FIGURE 6.—ROLLING 12-YEAR ESTIMATES OF LARGEST AUTOREGRESSIVE ROOT (WITH 90% CONFIDENCE INTERVAL)



Applying repeated iteration to the New Keynesian Phillips curve, equation (2), gives us an inflation process of the form

$$\pi_t = \gamma \sum_{k=0}^{\infty} \beta^k E_t y_{t+k}. \tag{9}$$

Combined with the output-gap process, this implies the following solution for inflation:

$$\pi_t = \frac{\gamma}{1 - \beta\phi} y_t. \tag{10}$$

In this case, the univariate processes for inflation and the output gap will be identical up to a scalar multiple. It is unlikely that this kind of example can fully explain our univariate results—an AR(4) regression for our output gap produces an OLS estimate of 0.76 for its persistence parameter, which is well below our inflation estimate of 0.96. However, this example shows that unless one conditions the inflation regression on appropriate driving variables, it is hard to make any direct link between the estimated value of ρ and the true effect on current inflation of its own lagged values.

It turns out, though, that the overall pattern of our results concerning the persistence parameter are little changed by the inclusion of a measure of the output gap, which we have constructed using a Hodrick-Prescott filter. Tables 4 and 5 report the full-sample OLS and grid-bootstrap estimates obtained from estimation of

$$\pi_t = \alpha + \rho\pi_{t-1} + \sum_{k=1}^{n} \psi_k \Delta\pi_{t-k} + \gamma y_t + \epsilon_t, \tag{11}$$

where y_t is the output gap. The first result worth noting is that, though admittedly crude, this measure of the output gap plays a statistically significant role in influencing inflation: The gap obtains a t-statistic of 4.5 in the GDP deflator regression, and 4.9 when added to the HICP specification. However, it has essentially no influence on estimates of the persistence parameter: The OLS and grid-bootstrap estimates show very little change from the univariate case.[16]

Table 6 reports the test results obtained from performing the unknown breakpoint tests on the GDP deflator specification including the output gap. As before, the hypothesis of no structural change in the persistence parameter cannot be rejected using the standard asymptotic distribution. Again, these tests suggest a break in the intercept that is significant at the 5% level, and conditioning on the estimated break (in

[16] In the grid-bootstrap estimation, the output-gap series was treated as a fixed regressor, with the same data series used across all of the bootstrap simulations.

TABLE 4.—OLS ESTIMATES OF OUTPUT GAP SPECIFICATION

	ρ	γ
GDP deflator	0.960 (0.039)	0.596 (0.132)
HICP	0.947 (0.043)	0.675 (0.137)

Standard errors in parentheses beside the coefficients.

TABLE 5.—GRID-BOOTSTRAP ESTIMATES FOR OUTPUT GAP SPECIFICATION

	Median Unbiased	90% Confidence Interval
GDP deflator	1.011	(0.932,1.055)
HICP	0.995	(0.917,1.051)

Note: Output gap treated as a fixed regressor in all bootstrap simulations.

1984:1), we obtain the lower estimate of $\rho = 0.78$. Again, though, adjusting for finite-sample bias by calculating *p*-values using bootstrap simulation methods casts considerable doubt on the statistical significance of the estimated intercept break: For example, the (unit-root-based) bootstrap estimate of the *p*-value for the intercept-break sup-*F* statistic is 22%. Finally, the pattern of results from our rolling regressions is almost completely unchanged by the addition of the output gap, so we do not report these calculations.[17]

VII. Interpreting the Results

One obvious interpretation of the results reported here is that they favor the simple backward-looking rule-of-thumb model of expectations over models that feature rational expectations. In particular, advocates of a rational-expectations approach would likely be surprised by the fact that the persistence parameter has remained stable despite the clear changes over time in monetary policy regimes, and by the fact that this parameter appears to have been high even through the 1990s and early 2000s, when the policy regime for the euro area could be argued to have had far more antiinflationary credibility. In addition, our general finding of parameter stability in a backward-looking model for inflation is consistent with Estrella and Fuhrer's (2003) conclusions, based on U.S. data, that such models tend to be more stable over time than models featuring rational expectations.

That said, we believe it is worth noting that it may be possible to reconcile our results with a popular class of models that feature both forward- and backward-looking expectations. Consider, for example, the recent work of Glenn Rudebusch (2005) on the empirical importance of the Lucas critique in New-Keynesian-style macroeconometric models. Rudebusch examines small multiequation models such as the following hybrid model for inflation, the output gap, and the short-term real interest rate that mixes both backward-looking and forward-looking rational expectations:

$$\pi_t = (1 - \mu_\pi)\pi_{t-1} + \mu_\pi E_t \pi_{t+1} + \alpha_y y_{t-1} + \epsilon_t, \qquad (12)$$

$$y_t = \beta_y[(1 - \mu_y) y_{t-1} + \mu_y E_t y_{t+1}] - \beta_r r_t + \eta_t, \qquad (13)$$

$$r_t = (1 - \mu_r)(i_{t-1} - \pi_{t-1}) + \mu_r(i_t - E_t \pi_{t+1}). \qquad (14)$$

[17] See the working-paper version of our paper, O'Reilly and Whelan (2004), for a chart of these rolling regression estimates.

Rudebusch solves for the reduced-form time series representation implied by this model. Perhaps surprisingly, he finds that the persistence parameter in inflation regressions varies little across a realistic range of values for the monetary policy rule. For example, for a highly forward-looking specification in which the weights on the expectational terms are all 0.8 or above, Rudebusch finds that switching from the estimated pre-Volcker policy rule to the post-Volcker rule (as estimated by Clarida, Galí, and Gertler, 2000) produces a change in the estimated persistence parameter from 0.32 to 0.23. The same shift in a more backward-looking model (in which the weights on expected inflation and output are lower than 0.3) leads to a change in the estimated persistence parameter from 0.99 to 0.95.

These calculations show that the joint existence of rational expectations and an inflation-targeting policy rule is not, on its own, a sufficient condition to imply fast convergence to an average value for inflation. Nor do changes in the monetary policy rule necessarily eliminate the usefulness of the full-sample reduced-form estimates for forecasting. Finally, although these calculations show that our finding of a high and stable value of ρ in reduced-form econometric equations can be reconciled with models featuring some role for rational expectations, they cannot be reconciled with models that feature *only* forward-looking rational expectations, such as the New Keynesian Phillips curve. Thus, our estimates are consistent with previous work by Rudd and Whelan (2005a, 2005b), who argue that the important role for lagged inflation terms in U.S. regressions cannot be reconciled with the pure New Keynesian model.

VIII. Conclusions

We have presented evidence on the stability over time of some simple reduced-form Phillips-curve equations for in-

TABLE 6.—STRUCTURAL-CHANGE TESTS FOR OUTPUT-GAP SPECIFICATION

	ρ-Break	Intercept Break
Sup-*F* (Andrews-Quandt) Tests		
Test values	6.72	8.65
Asymptotic *p*-values	0.12	0.05
Bootstrapped *p*-values ($\rho = 0.96$)	0.26	0.19
Bootstrapped *p*-values ($\rho = 1.00$)	0.28	0.22
Exp-*F* (Andrews-Ploberger) Tests		
Test values	1.88	2.42
Asymptotic *p*-values	0.06	0.03
Bootstrapped *p*-values ($\rho = 0.96$)	0.15	0.13
Bootstrapped *p*-values ($\rho = 1.00$)	0.17	0.15

Note: Bootstrapped *p*-values are based on simulating the estimated OLS process (the $\rho = 0.96$ case) or the estimated process with $\rho = 1$ imposed; shocks for the simulated processes were based on drawing from the estimated residuals.

flation in the euro area. Although large shifts in reduced-form coefficient estimates may have been expected as a response of rational agents to the sequence of shifts in monetary policy regimes that have taken place in the euro area since 1970, the overall message that we take away from our results is one of surprising stability in these coefficients. In particular, the important persistence parameter, which plays a crucial role in describing the impulse response patterns from inflationary shocks, appears to have been quite stable over the post-1970 period.

Our paper adds to a recent literature that has cast some doubt on the empirical relevance of the Lucas critique of reduced-form models. However, it is important to point out that the evidence presented in this paper cannot be used to rule out the possibility of future structural changes in the euro-area inflation process. It may indeed be the case—now that a hard and credible EMU has arrived—that inflation will become anchored near its target value and that the lag effect documented here will cease to play an important role. However, our analysis suggests that there is little historically based empirical evidence for the idea that the persistence of inflation will alter dramatically in response to these institutional changes.

REFERENCES

Andrews, Donald, "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica,* 61 (1993), 821–856.

Andrews, Donald, and Hong-Yuan Chen, "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics* 12 (1994), 187–204.

Andrews, Donald, and Werner Ploberger, "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative," *Econometrica* 61 (1994), 1383–1414.

Blanchard, Olivier, *Macroeconomics,* 2nd ed. (Prentice-Hall, 2000).

Clarida, Richard, Jordi Galí, and Mark Gertler, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature* 37 (1999), 1661–1707.

Clarida, Richard, Jordi Galí, and Mark Gertler, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics* 115 (2000), 147–180.

Cogley, Timothy, and Thomas Sargent, "Evolving Post-World War II Inflation Dynamics," in *NBER Macroeconomics Annual,* Vol. 16 (2002).

Diebold, Frank, and Celia Chen, "Testing Structural Stability with Endogenous Break Point: A Size Comparison of Analytic and Bootstrap Procedures," *Journal of Econometrics* 70 (1996), 221–241.

Estrella, Arturo, and Jeffrey Fuhrer, "Monetary Policy Shifts and the Stability of Monetary Policy Models," this REVIEW, 85:1 (2003), 94–104.

Fagan, Gabriel, Jerome Henry, and Ricardo Mestre, An Area-Wide Model (AWM) for the Euro Area, ECB Working paper no. 42 (2001).

Friedman, Milton, "The Role of Monetary Policy," *American Economic Review* 58 (1968), 1–17.

Hansen, Bruce, "Approximate Asymptotic *P* Values for Structural Change Tests," *Journal of Business and Economic Statistics* 15 (1997), 60–67.

Hansen, Bruce, "The Grid Bootstrap and the Autoregressive Model," this *REVIEW,* 81 (1999), 594–607.

Kieler, Mads, "Is Inflation Persistence Higher in the Euro Area than in the United States?" IMF country report no. 03/298 (2003).

Levin, Andrew, and Jeremy Piger, Is Inflation Persistence Intrinsic in Industrial Economies? Federal Reserve Bank of St. Louis working paper (2003).

Lucas, Robert, "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4 (1972a), 103–124.

Lucas, Robert, "Econometric Testing of the Natural Rate Hypothesis," in Otto Eckstein (Ed.), *The Econometrics of Price Determination Conference* (1972b), reprinted in Robert E. Lucas, Jr. (Ed.), *Studies in Business Cycle Theory* (Cambridge, MA: MIT Press, 1981).

Lucas, Robert, "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Series on Public Policy* 1 (1976) 19–46.

Marques, Carlos Robalo, "Inflation Persistence: Facts or Artefacts?' Banco de Portugal working paper (2003).

O'Reilly, Gerard, and Karl Whelan, "Has Euro-Area Inflation Persistence Changed over Time?" ECB working paper no. 335 (2004).

Perron, Pierre, "The Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Econometrica* 57 (1989), 1361–1401.

Pivetta, Frederic, and Ricardo Reis, "The Persistence of Inflation in the United States," Harvard University mimeograph (2003).

Quandt, Richard, "Tests of the Hypothesis That a Linear Regression Obeys Two Separate Regimes," *Journal of the American Statistical Association* 55 (1960), 324–330.

Roberts, John M., "New Keynesian Economics and the Phillips Curve," *Journal of Money, Credit, and Banking* 27 (1995), 975–984.

Rudd, Jeremy, and Karl Whelan, "New Tests of the New-Keynesian Phillips Curve," *Journal of Monetary Economics,* Vol. 52 (September 2005b).

Rudd, Jeremy, and Karl Whelan, "Does the Labor Share of Income Drive Inflation?" *Journal of Money, Credit, and Banking* 37 (2005a), 297–312.

Rudebusch, Glenn, "Assessing the Lucas Critique in Monetary Policy Models," *Journal of Money, Credit, and Banking* 37 (2005), 245–272.

Sargent, Thomas, "A Note on the 'Accelerationist' Controversy," *Journal of Money, Credit, and Banking* 3 (1971), 721–725.

Sargent, Thomas, *The Conquest of American Inflation* (Princeton: Princeton University Press, 1999).

Stock, James, "Confidence Intervals for the Largest Autoregressive Root in US Macroeconomic Time Series," *Journal of Monetary Economics* 28 (1991), 435–459.

Stock, James, "Comment on Cogley and Sargent," in *NBER Macroeconomics Annual,* Vol. 16 (2001).

Taylor, John B, "Monetary Policy Guidelines for Unemployment and Inflation Stability," in John Taylor and Robert Solow (Eds.), *Inflation, Unemployment, and Monetary Policy* (Cambridge, MA: MIT Press, 1998).