International Evidence on the Persistence of Inflation*

Luca Benati[†]
Bank of England

Abstract

Working with univariate autoregressions, we use tests for multiple structural breaks at unknown points in the sample, and the Stock-Watson (1996, 1998) time–varying parameters median-unbiased estimation methodology, to investigate the evolution and extent of inflation persistence over the post-WWII era for 41 quarterly inflation series from 20 OECD countries, plus the Eurozone. We construct confidence intervals for persistence estimates based on either Monte Carlo or bootstrapping procedures.

While, in most cases, estimates of persistence are characterised by a significant amount of uncertainty, results from structural break tests and, to a lesser extent, those based on the Stock-Watson methodology, suggest that in general inflation is *not* a highly persistent process: while it appears to be markedly persistent for some countries, during specific periods, empirical evidence suggests that, overall, high inflation persistence is not a robust feature of the data. Our results sound a cautionary note on taking high inflation persistence as a robust, established stylised fact that sticky-price DSGE models should necessarily replicate.

Keywords: Inflation; median-unbiased estimation; structural breaks; time-varying parameters; variance ratio; bootstrapping; frequency domain.

*Acknowledgements: I wish to thank Ken West for extremely helpful discussions, and Peter Andrews, Jushan Bai, Philip Hans Franses, Eric Ghysels, and Denise Osborn for useful suggestions. Previous versions of this work have been presented with a different title, 'Structural Breaks in Inflation Dynamics', at seminars at the 2003 North American Summer meetings of the Econometric Society, the 2003 meetings of the Society for Computational Economics, and at the Bank of England. Usual disclaimers apply. The views expressed in this paper are those of the author, and do not necessarily reflect those of the Bank of England.

†Bank of England, Threadneedle Street, London, EC2R 8AH. Email: Luca.Benati@bankofengland.co.uk. Phone: (+44)-020-7601-3573. Fax: (+44)-020-7601-4177.

1 Introduction

Inflation persistence has been, over the last decade, one of the most intensely investigated topics in the field of macroeconomics. Fuhrer and Moore (1995) first called the attention of the macroeconomics profession on the inability of 'standard'—i.e., Taylor (1980)—contracts to replicate the inflation persistence found in post-WWII U.S. data[1], and cited it as a major reason in favor of their alternative, relative contracting specification. Nelson (1998) argues that high inflation persistence is one of the two key stylised facts any sensible sticky-price dynamic stochastic general equilibrium (DSGE) model should be capable of replicating. More recently, Mankiw and Reis (2002), again, cite inflation persistence as a key motivation for their proposed 'sticky-information' model.

At a very general level, the notion that inflation is an *intrinsically* persistent process should however be seen, at the very least, with suspicion. First, on empirical grounds, Barsky (1987) first showed how the (alleged) high serial correlation typical of the post-WWII era was entirely absent under an alternative monetary regime, the Classical Gold Standard (1879-1914),[2] while Evans and Wachtel (1993), based on a Markow-switching model, showed how U.S. post-WWII inflation dynamics appears to have followed a two-state regime, displaying a unit root over the period 1968-82, but exhibiting much lower persistence either before 1968, or after 1982. Second, on strictly conceptual grounds, the stochastic properties of inflation cannot possibly be thought of as independent of the monetary regime in place over the sample period. A price-level targeting regime, for example, would make the price level (trend) stationary, thus causing inflation to be perfectly *negatively* serially correlated, in the sense that a 1 per cent positive shock to inflation today should necessarily be followed by a sequence of shocks into the infinite future exactly offsetting the initial jump in the price level. By the same token, as a simple matter of logic it is hard to believe that inflation may be a highly persistent process under an inflation-targeting regime in which the central bank pre-emptively and aggressively fights any deviation of inflation from target. The adoption, over the last decade, of inflation-targeting regimes in several developed and emerging economies raises therefore doubts on the notion that, today, inflation may be, in these countries, very highly persistent.

In recent years several papers have produced empirical evidence at odds with the notion of inflation as an intrinsically persistent process. Cogley and Sargent (2002) and Cogley and Sargent (2003), in particular, based on Bayesian random-coefficients VARs, detect evidence of a 'hump-shaped' evolution in the persistence of U.S. inflation over the post-WWII era, with persistence peaking around the time of the Great Inflation, but being essentially negligible both before mid-1960s, and over the most recent years. Levin and Piger (2003), building on a point first made by Perron

[1] An earlier paper, influential in establishing the conventional wisdom notion of inflation as a highly persistent process, was Nelson and Plosser (1982).

[2] On this issue, see also the pioneering work of Klein (1975).

(1989)—failure to control for possible structural breaks in the mean of a series spuriously increases its estimated extent of persistence—investigate inflation dynamics in 12 industrial countries over the period 1984-2001. For all series, they detect evidence of a single break in both the intercept and the innovation variance, but no evidence of a break in the autoregressive coefficients. Conditional on the identified breaks, all inflation series exhibit very little persistence.[3]

Working with univariate autoregressive representations, in this paper we use tests for multiple structural breaks at unknown points in the sample, and the Stock-Watson (1996, 1998) time-varying parameters median-unbiased estimation methodology, to investigate the evolution and extent of inflation persistence over the post-WWII era sfor 41 quarterly inflation series from 20 OECD countries, plus the Eurozone. We construct confidence intervals for persistence estimates based on either Monte Carlo or bootstrapping procedures. Our main findings may be summarised as follows. First, based on either methodology—and especially on Stock and Watson's—persistence estimates are characterised, in general, by a significant amount of uncertainty, to the point that it is often impossible to make strong statements concerning the precise extent of persistence for a specific series and/or sample period. Second, in spite of this, results from break tests and, to a lesser extent (due to the large econometric uncertainty), those based on Stock and Watson's methodology clearly show that, for some series and/or sample periods inflation has indeed been all but persistent, thus suggesting that, overall, high inflation persistence is *not* a robust feature of the data. Finally, even without allowing for any kind of time-variation in the data-generation process for inflation, results based on the Hansen (1999) 'grid bootstrap' procedure show that, for a few countries and series—the Netherlands, based on either the GDP deflator or the CPI; Korea, New Zealand, and Norway based on the GDP deflator; and Austria and Sweden based on the CPI—inflation has been all but persistent. Although our investigation must necessarily be regarded as preliminary, and our conclusions concerning the presence or absence of high persistence for a specific series and/or sample period as tentative, nonetheless this paper sounds a cautionary note on taking high inflation persistence as a *robust, established* stylised fact that sticky-price DSGE models should *necessarily* replicate.

The paper is organised as follows. The next section describes our dataset. In section 3 we present parametric measures of persistence based on univariate AR(p) representations for inflation series, both allowing and without allowing for time-variation in the underlying data generation process. Section 4 concludes.

[3] See also Kim, Nelson, and Piger (2004), Brainard and Perry (2000), and Benati (2004) for the U.S.; Ravenna (2000) for Canada; and Benati (2003a, 2003b) and Cogley, Morozov, and Sargent (2003) for the U.K..

2 The Data

Our dataset contains 41 quarterly inflation series from 20 OECD countries, plus the eurozone.[4] The respective sample periods are reported in Table 1. All the series are from the are from IMF's *International Financial Statistics* CD-ROM, with the following exceptions. The U.K. GDP and personal consumption expenditure deflators are from the *Office for National Statistics.* The U.S. GDP and personal consumption expenditure deflators are from Table 1.1.4. of the NIPA ('Price Indexes for Gross Domestic Product'). The eurozone's synthetic Harmonised Index of Consumer Prices (HICP), and the synthetic GDP and personal consumption expenditure deflators, are from the ECB's database. For Italy, the CPI excluding tobacco items is from *ISTAT*, and is available from January 1947 to September 2002, but we start the sample in January 1948 to prevent our results from being distorted by the extremely high inflation and subsequent stabilisation of 1947. Japan's CPI ('General Index Excluding Imputed Rent') is from Japan's *Statistics Bureau*'s website. The series is available from August 1946 to December 2002, but in what follows we focus on the period from January 1949 to March 1997, as, first, the period between the end of the war and the end of 1948 is characterised by a remarkable and anomalous volatility associated with the immediate aftermath of WWII; and second, the series has a discontinuity in April 1997. The CPI for Belgium is from the Belgian central bank. The CPIs for New Zealand, Sweden, Canada, France, Switzerland, the Netherlands, Austria, and the U.S. are from, respectively, *Statistics New Zealand*, *Statistics Sweden*, *Statistics Canada*, *INSEE*, *Ufficio Federale di Statistica*, *CBS* (the Dutch statistical office), *Statistik Austria*, and *U.S. Department of Labor, Bureau of Labor Statistics.* We also have results based on a quarterly CPI series for the whole of Germany, available for the period 1950:1-2002:2 (the series is from the *Bundesbank*'s monthly bullettin). While these results are available upon request, we have chosen not to report them as we have been unable to exactly determine how the issue of the reunification had been taken care of in the construction of the series.

Only the GDP and PCE deflators for the U.K., the U.S., and the eurozone, and the CPIs for Belgium, the Netherlands, the U.S., and Japan, are available on a seasonally adjusted basis. In all other cases, the original seasonally unadjusted data have been seasonally adjusted via the ARIMA X-12 procedure. While another possibility would have been to use seasonal dummies, we have preferred to seasonally adjust the data via a standardised procedure before performing the empirical analysis as, for many series, there are clear indications of changes in the seasonal pattern over the sample period (in several cases, the changes are so marked as to be detectable

[4] We wish to thank Jerome Henry of the ECB for kindly providing data for the eurozone; and Alberto Baffigi of *Banca d'Italia*, Graham Howard of the *Reserve Bank of New Zealand*, Peter Stadler of the Swiss central bank, Raf Wouters of the *National Bank of Belgium*, Cees Ullersma of the Dutch central bank, and Fabio Rumler of the Austrian central bank for providing CPI series for their respective countries.

with a simple Andrews-Ploberger test on the seasonal dummies based on univariate AR(p) representations for the inflation series; these results are available upon request). This, unfortunately, creates several problems, especially for the structural break tests of section 3.2.1. Including seasonal dummies in regressions based on the seasonally unadjusted series, and performing joint break tests for all the parameters, including the seasonal dummies, runs the risk of identifying breaks uniquely driven by changes in the seasonal pattern. On the other hand, running the same regressions, and testing only for breaks in the intercept and the AR coefficients presents the well-known[5] drawback of a diminished power of the test. Given the impossibility of finding an acceptable solution within this framework, we have therefore opted for the alternative of seasonally adjusting the series before performing the empirical analysis. As a justification for such an approach, it is important to keep in mind that the vast majority of macroeconomic time series do originally contain seasonal components, and that the seasonally adjusted data economists routinely work with represent the final product of seasonal adjustment procedures carried out inside statistical agencies, most of the times by means of some kind of filtering algorithm like ARIMA X-12. Although our approach is less-than-ideal, we believe it represent the most reasonable choice.

All the series are originally available at the quarterly frequency, with the exception of the CPIs for Canada, Italy, Switzerland, Belgium, the Netherlands, Austria, the U.S., and Japan. In all these cases, the original monthly series have been converted to the quarterly frequency by taking the last observation from each quarter.

Finally, in constrast to, e.g., Levin and Piger (2003), we do not use dummies to control for specific one-off events like the the Nixon price controls in the U.S. in the 1970s, or the introduction of the 'community charge' in the U.K. in April 1990. The key reason for doing so is to make our results exactly comparable to those found in the vast majority of the existing literature which, likewise, does not control for one-off events—see, e.g., Fuhrer and Moore (1995), Nelson (1998), Mankiw and Reis (2002), Evans and Wachtel (1993), and Cogley and Sargent (2002,2003).

3 Measuring persistence without allowing for time-variation

In this section we present parametric measures of persistence based on univariate AR(p) representations, without allowing for time-variation in the underlying data-generation process (henceforth, DGP). For each inflation series[6] we estimate via OLS the following AR(p) model

$$\pi_t = \mu + \phi_1 \pi_{t-1} + \phi_2 \pi_{t-2} + ... + \phi_p \pi_{t-p} + u_t \tag{1}$$

[5] See, e.g., Hansen (1992).

[6] We compute inflation as the non-annualised quarter-on-quarter rate of change of the relevant price index.

selecting the lag order based on the Schwartz information criterion, for a maximum possible number of lags P=6. Table 2 reports, for each series, the median-unbiased estimate of our preferred measure of persistence—which, following Andrews and Chen (1994), we take it to be the sum of the autoregressive coefficients[7]—computed via the Hansen (1999) 'grid bootstrap' procedure, together with the 90%-coverage confidence interval. Specifically, following Hansen (1999, section III.A) we recast (1) into the augmented Dickey-Fuller form

$$\pi_t = \mu + \rho\pi_{t-1} + \gamma_1\Delta\pi_{t-1} + ... + \Delta\gamma_{p-1}\pi_{t-(p-1)} + u_t \tag{2}$$

—where ρ is defined as the sum of the AR coefficients in (1)—and we simulate the sampling distribution of the t-statistic t=($\hat{\rho}$-ρ)/$\hat{S}(\hat{\rho})$, where $\hat{\rho}$ is the OLS estimate of ρ, and $\hat{S}(\hat{\rho})$ is its estimated standard error, over a grid of possible values [$\hat{\rho}$-3$\hat{S}(\hat{\rho})$; $\hat{\rho}$+3$\hat{S}(\hat{\rho})$], with step increments equal to 0.01. For each of the possible values in the grid, we consider 999 replications. Both the median-unbiased estimates of ρ and the 90% confidence intervals reported in Table 2 are based on the bootstrapped distribution of the t-statistic.

Several findings are clearly apparent from the table. Consistent with a vast literature, and with the Hansen 'grid bootstrap' estimates of ρ without allowing for structural breaks found in Levin and Piger (2003, section 2, table 1 and figure 2), many series indeed exhibit very high persistence. For the Eurozone, for example, we estimate unit root processes based on any of the three price indices[8], while for France, Italy, and Portugal we estimate unit root processes based on the GDP deflators, while we cannot reject the null of a unit root at the 90% level based on the CPIs. Overall, the null of a unit root cannot be rejected for 20 series out of 41 (48.8%), while the upper limits of the 90% confidence intervals are smaller than 0.95 for 17 series out of 41; smaller than 0.90 and 0.80 for 12 and, respectively, 7 series; and smaller than 0.70—corresponding to a half-life of reduced-form shocks of just 2 quarters—for 6 series out of 41.

Maybe not surprisingly, however, given our use of longer sample periods, our results detect, overall, less persistence than Levin and Piger, who cannot reject a unit root at the 90% level for 30 series out of 48 (62.5%), and for whom the upper limit of the 90% confidence interval is smaller than 0.95, 0.90, and 0.85 for only 7, 4, and, respectively 1 series. Further, our results clearly indicate that, for some individual price series, and even for some countries, inflation is all but very highly persistent. This is the case of the Netherlands based on either the GDP deflator

[7] As shown by Andrews and Chen (1994), the sum of the autoregressive coefficients maps one-to-one into two alternative measures of persistence, the cumulative impulse-response function to a one-time innovation and the spectrum at the frequency zero. Andrews and Chen (1994) also contain an extensive discussion of why an alternative measure favored, e.g., by Stock (1991), the largest autoregressive root, may provide a misleading indication of the true extent of persistence of the series depending on the specific values taken by the other autoregressive roots.

[8] Our results are therefore consistent with those of O'Reilly and Whelan (2004) based on the GDP deflator.

or the CPI; of Korea, New Zealand, and Norway based on the GDP deflator; and of Austria and Sweden based on the CPI. Even without allowing for any kind of time-variation in the underlying DGP—in particular in the mean, the feature whose variation, if unrecognized, may give rise to an overstatement of the authentic extent of persistence—the picture emerging from Table 2 is therefore not that of inflation as a *uniformly* very highly persistent process.

4 Measuring persistence allowing for time-variation

Although the Hansen 'grid bootstrap'median-unbiased estimates of the previuous section represent a useful starting point for our analysis, there are several reasons for contemplating the possibility that the data generation process for inflation may have not remained stable over time. First, on empirical grounds, the work of Stock and Watson (1996) has documented widespread instability in many macroeconomic time series, including inflation ones, while Barsky (1987), Evans and Wachtel (1993), Cogley and Sargent (2002, 2003), and Levin and Piger (2003) have documented time-variation in the stochastic properties of inflation over the last several decades. Second, on strictly conceptual grounds, the Lucas critique suggests that the marked shifts in the conduct of monetary policy which have characterised many countries in our sample over the last several decades—with the collapse of Bretton Woods being followed, for several of them, by a period with no clearly defined nominal anchor; and the introduction, over the last decade, of inflation-targeting regimes in Australia, Canada, New Zealand, Sweden, and the United Kingdom—should reasonably be expected to have produced changes in the stochastic properties of inflation.

For our purposes, the possibility of time-variation in the DGP for inflation has two main implications. First, the extent of inflation persistence may have changed over time, so that 'lumping together' different sample periods, with possibly different extents of persistence may provide a distorted picture. Second, to the extent that changes in the DGP for inflation have involved shifts in its equilibrium level, Perron's (1989) analysis suggests that this will automatically exaggerate the true extent of persistence. We therefore proceed to re-examine the issue of persistence allowing for two possible types of time-variation in univariate AR(p) representations, multiple structural breaks at unknown points in the sample, and random-walk time-varying parameters.

4.1 Tests for multiple structural breaks at unknown points in the sample

In this section we consider tests for multiple structural breaks at unknown points in the sample in either the sum of the AR coefficients, or both the intercept and the sum of the AR coefficients, in equation (1). In spite of our exclusive focus on persistence, there are several reasons for also considering tests for joint breaks in both features. First, and crucially, tests for breaks in individual features (i.e., tests for partial structural change) may have a remarkably low power—on this see, for example, Hansen (1992)—as critical values are derived under the maintained hypothesis that the features that are not being tested do not break, an assumption that, in practice, may not be satisfied.[9] While a rejection of stability in the sum of the AR coefficients should therefore be considered a very strong result, a failure to reject would not bear, in general, strong implications. Second, the classic Perron (1989) point—shifts in the conditional mean of a series, if not controlled for, may spuriously increase its estimated extent of persistence—imply that the features whose breaks are relevant for our purposes are not only the sum of the AR coefficients, but also the mean of the series: testing for joint breaks in both the intercept and the sum of the AR coefficients appears therefore as the most logical thing to do.

Our methodology combines the Andrews and Ploberger (1994) *exp*-Wald statistic and the Bai (1997a) method of estimating multiple breaks sequentially, one at a time.[10] There are several reasons for preferring this approach to the alternative Bai and Perron (1998, 2003) one. First, and crucially, the Bai-Perron methodology—as outlined in Bai and Perron (1998) and Bai and Perron (2003), and implemented in a *Gauss* code downloadable from Pierre Perron's web page—only allows to perform tests on the intercept and the AR coefficients *considered as a whole*, instead of the intercept and the *sum* of the AR coefficients, the key object of interest for us. Second, the Bai-Perron methodology is based on *sup*-type statistics, which, as discussed at length by Andrews and Ploberger (1994), are in general inferior to *exp*-type ones.[11]

For each inflation series we estimate (1) via OLS, and we start by testing for a structural break at an unknown point in the sample either in the sum of the AR coefficients, or in the intercept and the sum of the AR coefficients, based on the Andrews-Ploberger's (1994) exponential Wald statistic, imposing 15% symmetric trimming. We bootstrap the critical values as in Diebold and Chen (1996). The key reason for bootstrapping the critical values, instead of resorting to the asymptotic critical values tabulated in Andrews and Ploberger (1994), is that, as it is well known—see for example Diebold and Chen (1996)—the extent of size distortion tends to be quite high, first, in small samples; and second, for highly persistent processes. Given that, first,

[9] Quite obviously, booostrapping the critical values, as is done in the present case (see below), does not solve the problem, as bootstrapping is performed conditional on the assumption of no breaks in any of the model's features.

[10] As discussed in Bai (1997a), sequential estimation of the break dates, compared to the alternative simultaneous estimation, presents two key advantages. First, computational savings. Second, robustness to misspecification in the number of breaks.

[11] A third problem is that the Bai-Perron methodology is entirely based on asymptotic critical values, which suffer from a large extent of size distortion in small samples and for highly persistent processes (Monte Carlo evidence on this is available from the author upon request). This, however, could easily be fixed by bootstrapping the critical values.

for some series sample periods are relatively short;[12] and second, based on the results of the previous section, under the null of no breaks several series appear to be quite markedly persistent, basing test results on asymptotic critical values would run the risk of identifying spurious breaks, thus potentially distorting results for persistence. On the other hand, as shown by Diebold and Chen (1996), bootstrapped critical values perform in general well—in the sense of exhibiting little size distortion—even in small samples and in the case of highly persistent processes. We compute both test statistics, and boostrapped critical values, *via* a Newey and West (1987) correction for heteroskedasticity and/or autocorrelation (we set the number of lags in the Newey-West correction to 4).

If the null of no structural break is rejected, we estimate the break date by minimizing the residual sum of squares. The sample is then split in correspondence to the estimated break date, and the same procedure is repeated for each subsample. If the null of no structural break is not rejected for either subsample, the procedure is terminated. Otherwise, we estimate the new break date(s), we split the relevant subsample(s) in correspondence to the estimated break date(s), and we proceed to test for structural breaks for hierarchically obtained subsamples. The procedure goes on until, for each hierarchically obtained subsample, the null of no structural break is not rejected at the 10% level. After this preliminary, sequential estimation of the break dates, each break date is re-estimated according to Bai (1997a)'s repartition procedure.[13]

Tables 3 and 4 show the results. For each estimated break date we report bootstrapped p-values computed according to Diebold and Chen (1996), and 90% confidence intervals computed according to Bai (1997b), while for each sub-sample we report both the median-unbiased estimate of ρ and its 90% confidence interval, computed as in the previous section, based on the Hansen (1999) 'grid bootstrap' procedure. Several findings are clearly apparent from the tables. In particular,

(i) for several series and several sub-samples persistence estimates are characterised by a significant extent of uncertainty, to the point that it is not possible to make any statement concerning the presence of absence of persistence, as confidence intervals are compatible with both very high and very low persistence. Based on

[12] New Zealand's GDP deflator inflation series, for example, has only 66 observations, while the Japanese, Korean, and Sweden GDP deflator inflation series have less than 100 observations.

[13] More precisely, we re-estimate break dates according to the modification of the Bai (1997a) repartition procedure proposed by van Dijk, Osborn, and Sensier (2002)—specifically, each of the n estimated break dates is re-estimated conditional on the remaining n-1 break dates. In implementing the van Dijk, Osborn, and Sensier modification of the Bai procedure, we adopt the following iterative approach. We start by taking the first-stage estimated break dates as our initial conditions. Then, we re-estimate each break date conditional on the remaining n-1 break dates. These re-estimated break dates then become the initial conditions for the next iteration, and so on. The procedure is terminated when, from one iteration to the next, there is no difference in estimated break dates, so that we have reached a sort of 'econometric Nash equilibrium': each re-estimated break date is 'optimal'—in the sense of minimising the overall sum of the squared residuals—conditional on the remaining n-1 break dates.

tests for breaks in the sum of the AR coefficients for rates of inflation based on GDP deflators, for example, this is the case of France during the period 1970:2-1983:1; of Spain during the period 1978:1-1986:2; of Sweden during the period 1980:2-1991:4; and of Switzerland during the period 1991:3-2003:3.

(*ii*) In spite of this, there are many series and many sub-samples for which inflation is clearly *not* a very highly persistent process: based on tests for breaks in the sum of the AR coefficients for rates of inflation based on CPIs, this is the case, for example, of Austria and Norway over the entire sample periods; of Belgium over the period 1947:1-1963:4; of Canada over the sub-periods 1947:1-1966:3 and 1991:2-2002:2; of Denmark over the sub-period 1957:2-1985:1; of Finland over the sub-periods 1957:2-1971:4 and 1991:1-2003:4; of France over the sub-period 1982:3-2003:4; of Japan over the sub-period 1949:1-1964:2; of Sweden over the period 1947:1-1991:3; and of Switzerland over the period 1947:1-1967:4. Not surprisingly, allowing for structural breaks changes the overall picture quite markedly, compared with the one emerging from the previous sub-section, with a significant increase in the number of series exhibiting low persistence over at least one sub-period.

(*iii*) This does *not* imply, however, that high persistence has disappeared from the picture: not only it is still not possible to reject the null of a unit root at the 90% level for a significant number of series and sub-samples, but for several of them empirical evidence clearly points towards very high persistence. Based on tests for joint breaks in the intercept and the sum of the AR coefficients, this is the case of Italy post-1981:4, Portugal post-1983:3, of the U.K. post-1968:2, and the United States over the period 1958:3-1981:1 based on GDP deflators; and based on the CPIs, of fBelgium post-1964:1, of Canada over the period 1966:4-1991:1, of Finland post-1972:1, of Italy over the entire sample period, of Korea post-1976:4, of the Netherlands post-1970:3, of New Zealand over the period 1959:2-1986:4.

While our results are at odds with the traditional Fuhrer and Moore (1995)-Nelson (1998)-Mankiw and Reis (2002) view of inflation as a very highly persistent process, they are also not entirely compatible with Levin and Piger (2003) who, based on four price indices for twelve industrial countries over the period 1984-2001, detect evidence of a single break in the mean, but no evidence of a break in the sum of the autoregressive coefficients: as table 3 shows, for some series we do detect indeed, post-1984, evidence of a break in the sum of the AR coefficients. There are several explanations for these different results. First, Levin and Piger test for breaks in the sum of the AR coefficients treating as known—i.e., imposing—the previously identified breaks in the intercept, while our test is predicated on the assumption of no breaks in the intercept. Second, they use a stricter significance level, 95%. Even eliminating the post-1984 break dates with p-values greater than 5%, however, there are still some breaks left. Third, they use Andrews *sup*-Wald statistic, which is well-known—see, e.g., Andrews and Ploberger (1994)—for possessing lower power against local alternatives compared with the exponential Wald one. Fourth, and we suspect that this may be a key explanation, based on our own experience results from structural break tests are in general not robust to marked changes in the length of the sample period, so that (say) testing for a break in Sweden's CPI inflation over the post-1984 period may well produce different results than the same test performed, as we have done, over the period 1947:1-2001:3.[14]

4.2 Stock and Watson's (1996, 1998) time–varying parameters median-unbiased estimation methodology

In this section we present results based on the Stock and Watson (1996) and Stock and Watson (1998) time-varying parameters median-unbiased estimation methodology applied to model (1). With a single exception discussed below, concerning the issue of how to tackle the possible presence of heteroskedasticity in the data, we closely follow Stock and Watson (1996).

Letting θ_t=[μ_t, $\phi_{1,t}$, ..., $\phi_{p,t}$]$'$ and z_t=[1, $\pi_{t-1,t}$, ..., $\pi_{t-p,t}$]$'$, the time-varying parameters version of (1) is given by:

$$\pi_t = \theta_t' z_t + u_t \tag{3}$$

$$\theta_t = \theta_{t-1} + \eta_t \tag{4}$$

with η_t iid $N(0_{p+1},\ \lambda^2\sigma^2 Q)$, with 0_{p+1} being a (p+1)-dimensional vector of zeros; σ^2 being the variance of u_t; Q being a covariance matrix; and $E[\eta_t u_t]$=0. Following Nyblom (1989) and Stock and Watson (1996, 1998), we set Q=[$E(z_t z_t')$]$^{-1}$. Under such a normalisation, [$E(z_t z_t')$]$^{-1/2} z_t$ evolves according to a (p+1)-dimensional standard random walk, with λ^2 being the ratio between the variance of each 'transformed innovation' and the variance of u_t. [here stress the fact that λ is actually equal to tau/T]

Our point of departure are the Hansen (1999) grid bootstrap median-unbiased estimates of ρ of section 3.1. (Given that, as discussed in Stock and Watson (1998), for the TVP-MUB methodology to be applicable z_t ought not be integrated of order 1 or higher, the following procedure is only applied to those series for which $\hat{\rho}$<1 in table 2.) Conditional on the grid bootstrap estimate of ρ we estimate model (1); we compute the residuals, $\hat{u}_t$, and the estimate of the innovation variance, $\hat{\sigma}^2$; and we perform a Nyblom (1989) test of the null that λ=0 against the alternative that $\lambda \neq$0, based on the test statistic

$$\hat{L} = T^{-2} \sum_{t=1}^{T} \hat{S}_t' \hat{V}_t^{-1} \hat{S} \tag{5}$$

where

$$\hat{S}_t = \sum_{s=1}^{t} z_s \hat{u}_s, \tag{6}$$

[14] [check this conjecture by redoing all the tests over the post-1984 period]

and where

$$\hat{V}_t = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2 z_t z_t', \tag{7}$$

which is robust to the possible presence of heteroskedasticity—see Hansen (1990) and Stock and Watson (1996). Finally, we estimate the matrix Q as in Stock and Watson (1996) as

$$\hat{Q} = \left[T^{-1} \sum_{t=1}^{T} z_t z_t' \right]^{-1}. \tag{8}$$

We start by considering a 20-point grid of values for λ over the interval [0, 0.03], let's call it Λ. For each $\lambda_j \in \Lambda$ we compute the corresponding estimate of the covariance matrix of η_t as $\hat{Q}_j = \lambda_j^2 \hat{\sigma}^2 \hat{Q}$, and conditional on $\hat{Q}_j$ we simulate model (3)-(4) 10,000 times as in Stock and Watson (1996, section 2.4), with the only difference that here the pseudo innovations are generated by bootstrapping the residuals computed conditional on the Hansen grid bootstrap estimate of ρ (the $\hat{u}_t$),[15] instead of drawing from pseudo random *iid* $N(0,\ \hat{\sigma}^2)$. (For λ_j=0, we just bootstrap model (1) based on the OLS estimates conditional on the grid bootstrap estimate of ρ.) The key reason for bootstrapping the residuals, instead of just drawing from pseudo random *iid* $N(0,\ \hat{\sigma}^2)$ as in Stock and Watson (1996), is to take into account of the possible presence of heteroskedasticity in the data, which, if not controlled for, might distort the construction of the empirical distribution of the Nyblom test statistic. For each simulation, we compute a Nyblom (1989) test based on (5), thus building up its empirical distribution conditional on λ_j. Based on the empirical distributions of the $\hat{L}_j$'s we then compute the median-unbiased estimate of λ as that particular λ_j which is closest to the $\hat{L}$ we previously computed based on the actual data[16] (if $\hat{L}$ is greater than the median of the distribution of $\hat{L}_j$ conditional on λ_j=0.03, we set $\hat{\lambda}$=0.0316, adding one more 'step' to the grid). Finally, we compute the p-value for $\hat{L}$ based on the distribution of $\hat{L}_j$ conditional on λ_j=0. Table 5 reports, for each series, the Nyblom test statistic, the corresponding p-value, and the median-unbiased estimate of λ. At the 10% level, we detect evidence of random-walk time-variation for 18 series out of 41.

The next step is to compute time-varying estimates of ρ_t and, crucially, confidence bands around the estimates. In order to do that taking into account not only of filter, but also of parameter uncertainty, we need first of all to deconvolute the probability density function of $\hat{\lambda}$,[17] which we do *via* the following procedure. To fix ideas, let's start by considering the construction of a (1-α)% confidence interval for $\hat{\lambda}$,

[15] A comparatively minor problem with this is that the $\hat{u}_t$ are in general not the same as the residuals produced by model (3)-(4) conditional on λ>0. Although the problem might easily be tackled, we regard it as most likely insignificant for our purposes.

[16] [On the other hand, we do not interpolate as this would create problems with the procedure for deconvoluting the PDF of lambda]

[17] The key reason for deconvoluting the PDF of $\hat{\lambda}$ instead of postulating normality, and computing

$[\hat{\lambda}^L_{(1-\alpha)}, \hat{\lambda}^U_{(1-\alpha)}]$, and let's assume, for the sake of simplicity, that λ_j and $\hat{\lambda}$ can take any value over $[0; \infty)$. Given the duality between hypothesis testing and the construction of confidence intervals, the (1-α)% confidence set for $\hat{\lambda}$ comprises all the value of λ_j that cannot be rejected based on a two-sided test at the α% level. Given that an increase in λ_j automatically shifts the PDF of $\hat{L}_j$ conditional on λ_j upwards, $\hat{\lambda}^L_{(1-\alpha)}$ and $\hat{\lambda}^U_{(1-\alpha)}$ are therefore such that

$$P\left(\hat{L}_j > \hat{L} \mid \lambda_j = \hat{\lambda}^L_{(1-\alpha)}\right) = \alpha/2 \tag{B1}$$

$$P\left(\hat{L}_j < \hat{L} \mid \lambda_j = \hat{\lambda}^L_{(1-\alpha)}\right) = \alpha/2 \tag{B2}$$

Let $\phi_{\hat{\lambda}}(\lambda_j)$ and $\Phi_{\hat{\lambda}}(\lambda_j)$ be the probability density function and, respectively, the cumulative probability density function of $\hat{\lambda}$, defined over the domain of λ_j. The fact that $[\hat{\lambda}^L_{(1-\alpha)}, \hat{\lambda}^U_{(1-\alpha)}]$ is a (1-α)% confidence interval automatically implies that (1-α)% of the probability mass of $\phi_{\hat{\lambda}}(\lambda_j)$ lies between $\hat{\lambda}^L_{(1-\alpha)}$ and $\hat{\lambda}^U_{(1-\alpha)}$. This in turn implies that $\Phi_{\hat{\lambda}}(\hat{\lambda}^L_{(1-\alpha)})$=$\alpha/2$ and $\Phi_{\hat{\lambda}}(\hat{\lambda}^U_{(1-\alpha)})$=1-$\alpha/2$. Given that this holds for any $0 \leq \alpha \leq 1$, we therefore have that

$$\Phi_{\hat{\lambda}}(\lambda_j) = P\left(\hat{L}_j > \hat{L} \mid \lambda_j\right) \tag{B3}$$

In this way, based on the Nyblom test statistic, $\hat{L}$, and on the simulated distributions of the $\hat{L}_j$'s conditional on the λ_j's in Λ, we obtain an estimate of the cumulative probability density function of $\hat{\lambda}$ over the grid Λ, let's call it $\hat{\Phi}_{\hat{\lambda}}(\lambda_j)$. Finally, we fit a logistic function to $\hat{\Phi}_{\hat{\lambda}}(\lambda_j)$ via non-linear least squares, we compute the implied estimate of $\phi_{\hat{\lambda}}(\lambda_j)$—call it $\hat{\phi}_{\hat{\lambda}}(\lambda_j)$—and we rescale its elements so that they sum to one. Figure 1 shows, for each of the series in table 5 for which the bootstrapped p-values are smaller than 0.1, the deconvoluted PDFs of $\hat{\lambda}$, together with the TVP-MUB estimate of λ reported in table 5 (the vertical red bar).

Before proceeding further, it is necessary to briefly discuss the main difference betwen the approach adopted herein and the one found in, e.g., Stock and Watson (1996), concerning how we tackle the possible presence of heteroskedasticity in the data. As stressed by Stock (2002) in his discussion of Cogley and Sargent (2002),[18] estimating time-varying parameters models without controlling for the possible presence of heteroskedasticity causes a systematic overestimation of the authentic extent of coefficients' drift, as the imposition of a constant covariance structure forces the time-varying parameters to 'pick up' part of the variation in the data originating from time-variation in the covariance. For our specific purposes this is a very serious

a standard error of the estimate based on (say) Berndt, Hall, Hall, and Hausman (1974), is that the 'pile-up' problem discussed, e.g., by Stock and Watson (1998) causes the PDF of λ to be, in general, distinctly non-normal.

[18] The same point applies, e.g., to Pivetta and Reis (2004)—see their specification of the covariance matrix V on page 8.

issue, as overestimation of the extent of coefficients' drift—and in particular of drift in the conditional mean of the process—would automatically imply an *underestimation of persistence* at each point in time. Controlling for the possible presence of heteroskedasticity is therefore, in the present context, of paramount importance. In what follows we adopt a solution along the lines of Boivin (2004). Specifically, for each of the series in table 5 for which the bootstrapped p-values are smaller than 0.1, we test for multiple structural breaks at unknown points in the sample in the innovation variance[19] in equation (1), based on the Andrews-Ploberger *exp*-Wald statistic and the Bai (1997a) method of estimating multiple breaks sequentially, one at a time.[20] The methodology is exactly the same as described in section 3.2.1, with the only difference that we bootstrap critical values via the Hansen (2000) 'fixed regressor' bootstrap procedure. Out of 18 series for which the bootstrapped p-value in table 5 is smaller than 0.1, we detect at least one volatility break in all cases except for Japan's GDP deflator (results are not reported here for reasons of space, but are available from the author upon request).[21] In order to establish notation, let $\hat{\Omega}_h$=[$\hat{\sigma}^2_{h,1}$, ..., $\hat{\sigma}^2_{h,k}$]$'$ be the vector of estimated sub-sample volatilities for series h, for the k identified sub-samples. We are now ready to simulate the distribution of the sum of the AR coefficients at each point in time *via* the following Monte Carlo procedure (in what follows, $MN(v, V)$ indicates a multivariate normal distribution with mean v and covariance matrix V, while T is the sample length).

for i=1:N

- Get a draw for λ based on the deconvoluted, discretised PDF for $\hat{\lambda}$. Call this draw $\tilde{\lambda}_i$ For each of the $\hat{\sigma}^2_{h,z}$ in $\hat{\Omega}_h$, get a draw $\tilde{\sigma}^2_{h,z}$=[(T-p-1)$\hat{\sigma}^2_{h,z}$]/χ^2_{T-p-1}, thus getting the vector $\tilde{\Omega}_h$ for the simulated volatilities for each of the identified sub-periods.
- Conditional on $\tilde{\lambda}_i$ and the $\tilde{\sigma}^2_{h,z}$, z=1, ..., k, compute the covariance matrices of η_t in (4) for each of the k identified sub-period, as $\tilde{\lambda}_i^2\tilde{\sigma}^2_{h,z}\hat{Q}$.
- Conditional on $\tilde{\sigma}^2_{h,z}$ and $\tilde{\lambda}_i^2\tilde{\sigma}^2_{h,z}\hat{Q}$, z=1, ..., k, run the Kalman filter and smoother for (3)-(4), thus getting two-sided estimates of the state vector and of its precision matrix at each t, $\xi_{t|T}$ and, respectively, $P_{t|T}$.

[19] Boivin (2004) estimates two different variance for the pre-Volcker and post-1979 periods.

[20] As stressed by Boivin (2004, footnote 16), the estimation of different variances for different sub-samples is indeed 'entirely consistent with the TVP specification, asymptotically', given the assumption of local-to-zero time variation.

[21] One problem with this approach is that, given the well-known low power of tests for partial structural change, it is likely that that the number of identified volatility breaks is lower than the true number, so that we may likely end up only partially addressing the issue of heteroskedasticity. Unfortunately, it is not clear at all how to tackle this problem,

- For each t from p+1 to T, draw from $MN(\xi_{t|T}, P_{t|T})$. Call this draw $\xi^i_{t|T}$. Based on $\xi^i_{t|T}$, compute the sum of the AR coefficients, $\rho^i_{t|T}$.

  end

Based on the distribution of the $\rho^i_{t|T}$, we then compute both a median estimate of ρ_t (the blue line in figures 3-6), and 90% confidence intervals around the median (the red lines in the same figures). We set N=10,000. Finally, based on a single pass of the Kalman filter and smoother conditional on $\hat{\lambda}$ we compute the corresponding time-varying estimate of ρ, which is reported in figures 2-3 as a black line.

Table 5 and figures 2-3 point towards both similarities and differences with the results from structural break tests. In particular,

(i) with the exception of Norway and the Netherlands based on GDP deflator inflation, persistence estimates appear, once again, to be characterised by a significant extent of uncertainty, to the point that for most series it is not possible to make precise statements concerning the presence or absence of high persistence. In particular, for all series uncertainty is so large to apparently suggest that it would not possible, for any of them, to reject the null of time-invariance. As stressed by Boivin (2004), however, such a conclusion would be incorrect, as the presence of time-variation is testified by the stability test results reported in table 5: so, although there is indeed a large uncertainty concerning the precise value taken by __ at each point in time, the Nyblom tests results clearly indicate that parameters drift is actually there.

(ii) In spite of such a large uncertainty, for a few countries and periods inflation clearly appears to be all but persistent. Based on the GDP deflator, this is the case for the Netherlands and Norway over the entire sample periods, and of Japan over the first part of the sample, while based on the CPI it is the case for Norway over most of the sample, and possibly for Sweden. Compared with the results from break tests, however, the evidence is much weaker, due to the much greater overall uncertainty.

(iii) In contrast with the results from break tests, for *no* series and/or period it is possible to state with reasonable confidence that inflation is very highly persistent. Although for many series and periods the upper 90% confidence band for ρ_t stretches close to 1, the extent of uncertainty is such that the very same estimates are compatible with low persistence.

Overall, results based on the Stock-Watson TVP-MUB methodology appear therefore as much more inconclusive than those from structural break tests, with the extent of uncertainty being so large as to make it essentially impossible to make strong statements about persistence—the possible exception being that, for a few series and/or sample periods, inflation appears to be all but persistent.

5 Conclusions

Motivated by the intensity of the attention paid by the macroeconomic profession to the issue of inflation persistence over the last decade, and by several authors' suggestion that the ability of sticky-price DSGE models to generate high persistence should be regarded as a crucial test of adequacy—see, e.g., Fuhrer and Moore (1995), Nelson (1998), and Mankiw and Reis (2002)—in this paper we have applied both tests for multiple breaks at unknown points in the sample, and the Stock-Watson (1996, 1998) time-varying parameters median-unbiased estimation methodology, to univariate AR(p) representations for inflation series, in order to investigate the evolution and extent of inflation persistence over the post-WWII era for 41 series from 20 OECD countries, plus the eurozone. Our main results may be summarised as follows.

(i) Based on either methodology—and especially on Stock and Watson's—persistence estimates are characterised, in general, by a significant amount of uncertainty, to the point that it is often impossible to make strong statements concerning the precise extent of persistence for a specific series and/or sample period.

(ii) In spite of this, results from break tests and, to a much lower extent (due to the large econometric uncertainty), those based on Stock and Watson's methodology clearly suggest that, for some series and/or sample periods inflation has indeed been all but persistent, thus suggesting that, overall, high inflation persistence is *not* a robust feature of the data.

(iii) Finally, it is important to stress how, even without allowing for any kind of time-variation in the data-generation process for inflation, results based on the Hansen (1999) 'grid bootstrap' procedure show that, for a few countries and series—the Netherlands, based on either the GDP deflator or the CPI; Korea, New Zealand, and Norway based on the GDP deflator; and Austria and Sweden based on the CPI—inflation has been all but persistent.

Although our investigation must necessarily be regarded as preliminary, and our conclusions concerning the presence or absence of high persistence for a specific series and/or sample period as tentative, nonetheless this paper sounds a cautionary note on taking high inflation persistence as a *robust, established* stylised fact that sticky-price DSGE models should *necessarily* replicate.

References

ANDREWS, D., AND W. CHEN (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12, 187–204.

ANDREWS, D. K., AND W. PLOBERGER (1994): "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62(6), 1383–1414.

BAI, J. (1997a): "Estimating Multiple Breaks One at a Time," *Econometric Theory*, 13, 315–352.

——— (1997b): "Estimation of a Change Point in Multiple Regression Models," *Review of Economics and Statistics*, 79, 551–563.

BAI, J., AND P. PERRON (1998): "Estimating and Testing Linear Models with Multiple Structural Changes.," *Econometrica*, 66(1), 47–78.

——— (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18, 1–22.

BARSKY, R. (1987): "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," *Journal of Monetary Economics*, 19, 3–24.

BENATI, L. (2004): "Investigating Inflation Persistence Across Monetary Regimes," *Bank of England Working Paper, forthcoming.*

BERNDT, E., B. HALL, R. HALL, AND J. HAUSMAN (1974): "Estimation and Inference in Non-Linear Structural Models," *Annals of Economic and Social Measurement*, 3-4, 653–665.

BOIVIN, J. (2004): "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," Columbia University, mimeo.

BRAINARD, W., AND G. PERRY (2000): "Making Policy in a Changing World," *in Economic Events, Ideas, and Policies: The 1960s and After, Perry,, G.L., and Tobin, J., eds., Brookings Institution Press.*

COGLEY, T., S. MOROZOV, AND T. J. SARGENT (2003): "Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System," University of California at Davis, Stanford University, and New York University, mimeo.

COGLEY, T., AND T. J. SARGENT (2002): "Evolving Post-WWII U.S. Inflation Dynamics," *in B. Bernanke and K. Rogoff, eds. (2002), NBER Macroeconomics Annuals 2001.*

——— (2003): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," University of California at Davis and New York University, mimeo.

Diebold, F. X., and C. Chen (1996): "Testing Structural Stability with Endogenous Breakpoint: A Size Comparison of Analytic and Bootstrap Procedures," *Journal of Econometrics*, 70, 221–241.

Evans, M., and P. Wachtel (1993): "Inflation Regimes and the Sources of Inflation Uncertainty," *Journal of Money, Credit, and Banking*, 25, 475–511.

Fuhrer, J., and G. Moore (1995): "Inflation Persistence," *Quarterly Journal of Economics*, 109, 127–159.

Hansen, B. (1990): "Lagrange Multiplier Tests for Parameter Instability in Non-Linear Models," available at http://www.ssc.wisc.edu/ bhansen/papers/old.htm.

Hansen, B. (1992): "Testing for Parameter Instability in Linear Models," *Journal of Policy Modelling*, 14, 517–533.

Hansen, B. (1999): "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics*, 81(4), 594–607.

Hansen, B. (2000): "Testing for Structural Change in Conditional Models," *Journal of Econometrics*, 97, 93–115.

Kim, C. J., C. Nelson, and J. Piger (2004): "The Less Volatile U.S. Economy: A Bayesian Investigation of Timing, Breadth, and Potential Explanations," *Journal of Business and Economic Statistics*, 22(1), 80–93.

Klein, B. (1975): "Our New Monetary Standard: Measurement and Effects of Price Uncertainty, 1880-1973," *Economic Inquiry*, 13, 461–484.

Levin, A. T., and J. Piger (2003): "Is Inflation Persistence Intrinsic in Industrial Economies?," Federal Reserve Bank of S. Louis, working paper.

Mankiw, N. G., and R. Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*.

Nelson, C. R., and C. I. Plosser (1982): "Trends and Random Walks in Macro-Economic Time Series: Some Evidence and Implications.," *Journal of Monetary Economics*, 10, 139–162.

Nelson, E. (1998): "Sluggish Inflation and Optimising Models of the Business Cycle," *Journal of Monetary Economics*, 42, 303–322.

Newey, W., and K. West (1987): "A Simple Positive-Semi-Definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Nyblom, J. (1989): "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84(405), 223–230.

O'Reilly, G., and K. Whelan (2004): "Has Inflation Persistence in the Euro Area Changed Over Time?," *European Central Bank Working Paper Series*, (??).

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

Pivetta, F., and R. Reis (2004): "The Persistence of Inflation in the United States," Harvard University, mimeo.

Ravenna, F. (2000): "The Impact of Inflation Targeting in Canada: A Structural Analysis," University of California S. Cruz, mimeo.

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

Stock, J. (2002): "Discussion," *in B. Bernanke and K. Rogoff, eds. (2002), NBER Macroeconomics Annuals 2001.*

Stock, J., and M. Watson (1996): "Evidence of Structural Instability in Macroeconomic Time Series Relations," *Journal of Business ands Economic Statistics*, 14(1), 11–30.

——— (1998): "Median-Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model," *Journal of the Americal Statistical Association*, 93(441), 349–358.

Taylor, J. (1980): "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 88, 1–22.

van Dijk, D., D. Osborn, and M. Sensier (2002): "Changes in the Variability of the Business Cycle in the G7 Countries," CGBCR Discussion Paper 016, University of Manchester.

Table 1 The sample periods			
	GDP deflator	CPI	PCE deflator
Australia	1959:4-2003:4	1957:2-2003:4	
Austria	1964:2-2003:4	1950:1-2002:2	
Belgium		1947:1-2002:3	
Canada	1957:2-2003:3	1947:1-2002:2	
Denmark		1957:2-2003:4	
Eurozone	1970:2-2002:4	1970:2-2002:4[#]	1970:2-2002:4
Finland	1970:2-2003:4	1957:2-2003:4	
France	1970:2-2003:4	1957:2-2003:4	
Ireland		1957:2-2003:4	
Italy	1970:2-2003:4	1948:1-2003:4	
Japan	1980:1-2003:4	1949:1-1997:1	
Korea	1980:1-2003:4	1970:2-2003:4	
Netherlands	1977:2-2003:4	1947:1-2002:3	
New Zealand	1987:3-2003:4	1947:1-2002:2	
Norway	1966:2-2003:4	1957:2-2003:4	
Portugal	1977:2-2003:4	1957:2-2003:4	
Spain	1970:2-2003:4	1957:2-2003:4	
Sweden	1980:2-2003:4	1947:1-2001:3	
Switzerland	1970:2-2003:3	1947:1-2003:4	
U.K.	1955:2-2003:4		1955:2-2003:4
U.S.	1947:2-2003:4	1947:2-2003:4	1947:2-2003:4

[#] HICP.

Table 2 Hansen (1999) 'grid-bootstrap' median-unbiased estimates of ρ, and 90% confidence intervals, without allowing for time-variation

	GDP deflator	CPI	PCE deflator
Australia	0.91 [0.79; 1.03]	0.90 [0.81; 1.01]	
Austria	0.94 [0.71; 1.07]	0.44 [0.29; 0.59]	
Belgium		0.70 [0.57; 0.84]	
Canada	0.89 [0.80; 1.00]	0.77 [0.68; 0.86]	
Denmark		0.78 [0.66; 0.90]	
Eurozone	1.01 [0.94; 1.04]	1.00 [0.92; 1.02]	1.00 [0.93; 1.03]
Finland	0.90 [0.70; 1.05]	0.87 [0.79; 0.95]	
France	1.01 [0.93; 1.05]	0.97 [0.89; 1.02]	
Ireland		0.91 [0.81; 1.02]	
Italy	1.00 [0.89; 1.04]	0.94 [0.86; 1.02]	
Japan	0.78 [0.64; 0.97]	0.67 [0.54; 0.81]	
Korea	0.46 [0.25; 0.66]	0.87 [0.74; 1.01]	
Netherlands	0.45 [0.29; 0.60]	0.52 [0.36; 0.68]	
New Zealand	0.05 [-0.34; 0.46]	0.89 [0.80; 0.99]	
Norway	0.04 [-0.09; 0.18]	0.73 [0.63; 0.84]	
Portugal	1.01 [0.90; 1.07]	0.94 [0.84; 1.03]	
Spain	0.90 [0.76; 1.03]	0.93 [0.84; 1.02]	
Sweden	0.72 [0.43; 1.04]	0.64 [0.53; 0.75]	
Switzerland	0.75 [0.6; 0.92]	0.78 [0.67; 0.88]	
U.K.	0.83 [0.72; 0.95]		0.93 [0.84; 1.02]
U.S.	0.86 [0.79; 0.94]	0.82 [0.73; 0.92]	0.86 [0.79; 0.93]

[#] HICP. For technical details, see section 2.1.

Table 3 Tests for multiple breaks at unknown points in the sample in the sum of the AR coefficients

Break dates and 90% confidence intervals	*exp*-Wald (p-value)	Sub-periods	$\hat{\rho}$ and 90% confidence intervals
	(a) GDP deflators		
	Austria		
1971:2 [1967:4; 1974:4]	6.17 (0.091)	1964:2-1971:1	1.64 [0.99; 2.48]
		1971:2-2003:4	0.85 [0.68; 1.03]
	France		
1983:2 [1982:2; 1984:2]	6.939 (0.069)	1970:2-1983:1	0.60 [0.21; 1.04]
		1983:2-2003:4	0.60 [0.37; 0.87]
	Netherlands		
1981:4 [1980:1; 1983:3]	5.516 (0.039)	1977:2-1981:3	0.45 [-0.11; 1.02]
		1981:4-2003:4	0.29 [0.11; 0.47]
	Spain		
1978:1 [1977:4; 1978:2]	58.251 (0.001)	1970:2-1977:4	0.89 [0.78; 1.01]
1986:3 [1985:2; 1987:4]	16.389 (0.048)	1978:1-1986:2	0.40 [-0.38; 1.07]
1992:2 [1991:1; 1993:3]	33.581 (0.01)	1986:3-1992:1	-0.74 [-2.02; 0.67]
		1992:2-2003:4	-0.18 [-0.66; 0.34]
	Sweden		
1992:1 [1991:1; 1993:1]	17.054 (0.041)	1980:2-1991:4	0.53 [-0.19; 1.09]
		1992:1-2003:4	-0.80 [-1.62; -0.02]
	Switzerland		
1991:3 [1988:3; 1994:3]	8.228 (0.04)	1970:2-1991:2	0.59 [0.34; 0.88]
		1991:3-2003:3	0.51 [0.09; 1.02]
	United Kingdom		
1963:2 [1962:2; 1964:2]	4.083 (0.079)	1955:2-1963:1	-0.43 [-1.39; 0.54]
		1963:2-2003:4	0.88 [0.77; 1.01]
	(b) CPIs		
	Austria		
1957:4 [1956:1; 1959:3]	7.886 (0.007)	1950:1-1957:3	-0.05 [-0.94; 0.82]
		1957:4-2002:2	0.73 [0.57; 0.94]
	Belgium		
1964:1 [1961:4; 1966:2]	34.148 (1.0E-3)	1947:1-1963:4	-0.02 [-0.50; 0.52]
		1964:1-2002:3	0.93 [0.80; 1.03]
	Canada		
1966:4 [1963:1; 1970:3]	7.852 (0.021)	1947:1-1966:3	0.49 [0.23; 0.74]
1991:2 [1990:4; 1991:4]	8.731 (0.036)	1966:4-1991:1	0.90 [0.75; 1.03]
		1991:2-2002:2	-0.03 [-0.64; 0.81]

For technical details, see section 2.2.1.

Table 3 (continued) Tests for multiple breaks at unknown points in the sample in the sum of the AR coefficients			
Break dates and 90% confidence intervals	*exp*-Wald (p-value)	Sub-periods	$\hat{\rho}$ and 90% confidence intervals
	(b) CPIs		
	Denmark		
1985:2 [1981:3; 1989:1]	6.363 (0.03)	1957:2-1985:1	0.63 [0.45; 0.83]
		1985:2-2003:4	0.72 [0.52; 0.97]
	Finland		
1972:1 [1968:3; 1975:3]	9.286 (0.008)	1957:2-1971:4	0.56 [0.33; 0.77]
1991:1 [1990:2; 1991:4]	7.388 (0.074)	1972:1-1990:4	0.99 [0.82; 1.03]
		1991:1-2003:4	0.27 [-0.07; 0.57]
	France		
1965:4 [1963:3; 1968:1]	6.128 (0.073)	1957:2-1965:3	0.48 [-0.69; 1.10]
1982:3 [1981:1; 1984:1]	8.987 (0.048)	1965:4-1982:2	0.98 [0.78; 1.05]
		1982:3-2003:4	0.64 [0.47; 0.82]
	Ireland		
1981:4 [1977:4; 1985:4]	5.575 (0.041)	1957:2-1981:3	0.99 [0.78; 1.05]
		1981:4-2003:4	0.84 [0.69; 1.01]
	Japan		
1964:3 [1961:3; 1967:3]	4.883 (0.071)	1980:1-1964:2	0.46 [0.08; 0.94]
1974:1 [1970:4; 1977:2]	8.269 (0.045)	1964:3-1973:4	1.01 [-0.42; 1.06]
		1974:1-2003:4	0.81 [0.62; 0.99]
	Korea		
1976:4 [1973:4; 1979:4]	10.907 (0.014)	1980:1-1976:3	0.69 [0.34; 1.02]
		1976:4-2003:4	0.94 [0.80; 1.04]
	New Zealand		
1986:4 [1982:3; 1991:1]	18.331 (1.0E-3)	1947:1-1986:3	0.91 [0.81; 1.02]
		1986:4-2002:2	0.76 [0.46; 1.04]
	Norway		
1970:1 [1966:3; 1973:3]	7.157 (0.02)	1957:2-1960:4	0.41 [0.11; 0.70]
1988:3 [1987:4; 1989:2]	3.571 (0.054)	1970:1-1988:2	0.64 [0.44; 0.88]
		1988:3-2003:4	0.02 [-0.26; 0.35]
	Sweden		
1991:4 [1986:3; 1997:1]	13.670 (2.0E-3)	1947:1-1991:3	0.59 [0.46; 0.72]
		1991:4-2001:3	0.67 [0.44; 1.01]
	Switzerland		
1968:1 [1965:4; 1970:2]	4.923 (0.029)	1947:1-1967:4	0.64 [0.47; 0.83]
		1968:1-2003:4	0.87 [0.73; 1.02]
For technical details, see section 2.2.1.			

Table 3 (continued) Tests for multiple breaks at unknown points in the sample in the sum of the AR coefficients

Break dates and 90% confidence intervals	*exp*-Wald (p-value)	Sub-periods	$\hat{\rho}$ and 90% confidence intervals
	(c) PCE deflators		
	United Kingdom		
1975:2 [1971:1; 1979:3]	14.207 (9.0E-3)	1955:2-1975:1	1.21 [1.05; 1.30]
		1975:2-2003:4	0.85 [0.73; 0.98]
	United States		
1967:1 [1964:4; 1969:2]	2.750 (0.021)	1947:2-1966:4	0.42 [0.18; 0.67]
		1967:1-2003:4	0.96 [0.88; 1.02]
For technical details, see section 2.2.1.			

Table 4 Tests for multiple breaks at unknown points in the sample in the intercept and the sum of the AR coefficients

Break dates and 90% confidence intervals	*exp*-Wald (p-value)	Sub-periods	$\hat{\rho}$ and 90% confidence intervals
	(a) GDP deflators		
	France		
1983:2 [1982:2; 1984:2]	33.282 (0.082)	1970:2-1983:1	0.60 [0.22; 1.03]
		1983:2-2003:4	0.61 [0.39; 0.87]
	Italy		
1981:4 [1981:1; 1982:3]	525.378 (0.004)	1970:2-1981:3	0.13 [-0.37; 0.66]
		1981:4-2003:4	0.90 [0.76; 1.03]
	Netherlands		
1986:4 [1983:4; 1989:4]	27.100 (0.02)	1977:2-1986:3	0.79 [0.48; 1.03]
		1986:4-2003:4	0.03 [-0.17; 0.23]
	Portugal		
1983:3 [1983:2; 1983:4]	101.893 (0.056)	1977:2-1983:2	-2.34 [-4.73; 0.72]
		1983:3-2003:4	0.91 [0.74; 1.04]
	Spain		
1980:1 [1979:3; 1980:3]	72.540 (0.030)	1970:2-1979:4	0.81 [0.63; 0.99][a]
1992:1 [1991:1; 1993:1]	49.521 (0.063)	1980:1-1992:1	0.29 [-0.20; 0.81]
		1992:1-2003:4	-0.12 [-0.63; 0.44]
	Sweden		
1992:1 [1991:1; 1993:1]	69.732 (0.087)	1980:2-1991:4	0.52 [-0.17; 1.09]
		1992:1-2003:4	-0.77 [-1.59; 0.00]
	Switzerland		
1975:1 [1974:4; 1975:2]	14.761 (0.083)	1970:2-1974:4	-0.80 [-2.05; 0.61]
		1975:1-2003:3	0.69 [0.48; 0.89]
	U.K.		
1968:2 [1967:3; 1969:1]	10.735 (0.070)	1955:2-1968:1	-0.27 [-0.81; 0.35]
		1968:2-2003:4	0.89 [0.77; 1.02]
	U.S.		
1958:3 [1954:3; 1962:3]	9.878 (0.031)	1947:2-1958:2	0.60 [0.32; 0.97]
1981:2 [1977:4; 1984:4]	15.233 (0.092)	1958:3-1981:1	1.00 [0.90; 1.03]
		1981:2-2003:4	0.77 [0.65; 0.91]
	(b) CPIs		
	Austria		
1957:4 [1956:1; 1959:3]	18.181 (0.013)	1950:1-1957:3	-0.07 [-0.92; 0.84]
		1957:4-2002:2	0.74 0.56; 0.92]
	Belgium		
1964:1 [1961:4; 1966:2]	39.632 (0.011)	1947:1-1963:4	-0.02 [-0.50; 0.47]
		1964:1-2002:3	0.93 [0.79; 1.03]

For technical details, see section 2.2.1. [a] Based on simple OLS, as the Hansen (1999) procedure produced unreliable results.

Table 4 (continued) Tests for multiple breaks at unknown points in the sample in the sample in the intercept and the sum of the AR coefficients

Break dates and 90% confidence intervals	*exp*-Wald (p-value)	Sub-periods	$\hat{\rho}$ and 90% confidence intervals
	(b) CPIs		
	Canada		
1966:4 [1963:1; 1970:3]	14.543 (0.031)	1947:1-1966:3	0.49 [0.24 ; 0.75]
1991:2 [1990:4; 1991:4]	19.220 (0.085)	1966:4-1991:1	0.90 [0.74; 1.03]
		1991:2-2002:2	-0.03 [-0.71; 0.80]
	Finland		
1972:1 [1969:1; 1975:1]	11.108 (0.059)	1957:2-1971:4	0.54 [0.32; 0.78]
		1972:1-2003:4	0.99 [0.90; 1.02]
	Italy		
1956:2 [1947:3; 1965:1]	15.842 (0.038)	1948:1-1956:1	1.00 [-0.17; 1.06]
1974:3 [1970:1; 1979:1]	17.447 (0.067)	1956:2-1974:2	1.10 [1.00; 1.35]
		1974:3-2003:4	0.99 [0.93; 1.03]
	Korea		
1976:4 [1973:4; 1979:4]	55.595 (0.025)	1970:2-1976:3	0.68 [0.37; 1.02]
		1976:4-2003:4	0.94 [0.80; 1.04]
	Netherlands		
1970:3 [1967:2; 1973:4]	5.360 (0.08)	1947:1-1970:2	0.22 [-0.08; 0.49]
		1970:3-2002:3	0.94 [0.82; 1.03]
	New Zealand		
1959:2 [1954:3; 1964:1]		1947:1-1959:1	0.75 [0.33; 1.06]
1987:1 [1984:4; 1989:2]	48.651 (0.007)	1959:2-1986:4	1.00 [0.85; 1.04]
		1987:1-2002:2	0.71 [0.42; 1.04]
	Norway		
1990:4 [1989:4; 1991:4]	5.638 (0.093)	1957:2-1990:3	0.68 [0.55; 0.84]
		1990:4-2003:4	-0.19 [-0.56; 0.17]
	Portugal		
1976:3 [1969:1; 1984:1]	73.706 (0.009)	1957:2-1976:2	0.87 [0.65; 1.05]
		1976:3-2003:4	0.97 [0.87; 1.03]
	Sweden		
1991:4 [1986:3; 1997:1]	9.974 (0.019)	1947:1-1991:3	0.59 [0.46; 0.72]
		1991:4-2001:3	0.68 [0.42; 1.02]
	(c) PCE deflators		
	U.K.		
1975:2 [1971:1; 1979:3]	25.428 (0.051)	1970:2-1975:1	1.21 [1.05; 1.37]
		1975:2-2002:4	0.85 [0.73; 0.99]

For technical details, see section 2.2.1.

Table 5 Results based on the Stock-Watson (1996, 1998) TVP-MUB methodology: Nyblom test statistics, bootstrapped *p*-values, and median-unbiased estimates of λ

	GDP deflator			CPI			PCE deflator		
	$\hat{L}$	p-value	$\hat{\lambda}$	$\hat{L}$	p-value	$\hat{\lambda}$	$\hat{L}$	p-value	$\hat{\lambda}$
Australia	0.595	0.770	0	0.559	0.580	~ 0			
Austria	1.624	0.104	0.017	2.037	0.020	0.032			
Belgium				1.656	0.014	0.025			
Canada	0.561	0.770	0	2.198	0.002	0.032			
Denmark				2.110	0.003	0.032			
Eurozone	NA	NA	NA	NA#	NA#	NA#	NA	NA	NA
Finland	1.214	0.207	0.016	1.434	0.003	0.032			
France	NA	NA	NA	1.908	0.037	0.025			
Ireland				0.747	0.299	0.011			
Italy	NA	NA	NA	1.120	0.401	0.006			
Japan	1.332	0.018	0.032	1.617	0.045	0.022			
Korea	1.051	0.188	0.032	0.888	0.156	0.022			
Netherlands	0.612	0.082	0.032	1.230	0.046	0.021			
New Zealand	0.740	0.729	0	0.811	0.644	0			
Norway	0.767	0.041	0.032	1.098	0.020	0.032			
Portugal	NA	NA	NA	1.111	0.288	0.009			
Spain	1.196	0.041	0.030	1.503	0.079	0.017			
Sweden	0.569	0.924	0	1.252	0.014	0.032			
Switzerland	1.153	0.052	0.032	1.269	0.032	0.021			
U.K.	0.854	0.202	0.014				0.938	0.306	0.009
U.S.	0.642	0.198	0.013	1.167	0.396	0.005	1.323	0.071	0.016

HICP. For technical details, see section 2.2.2.

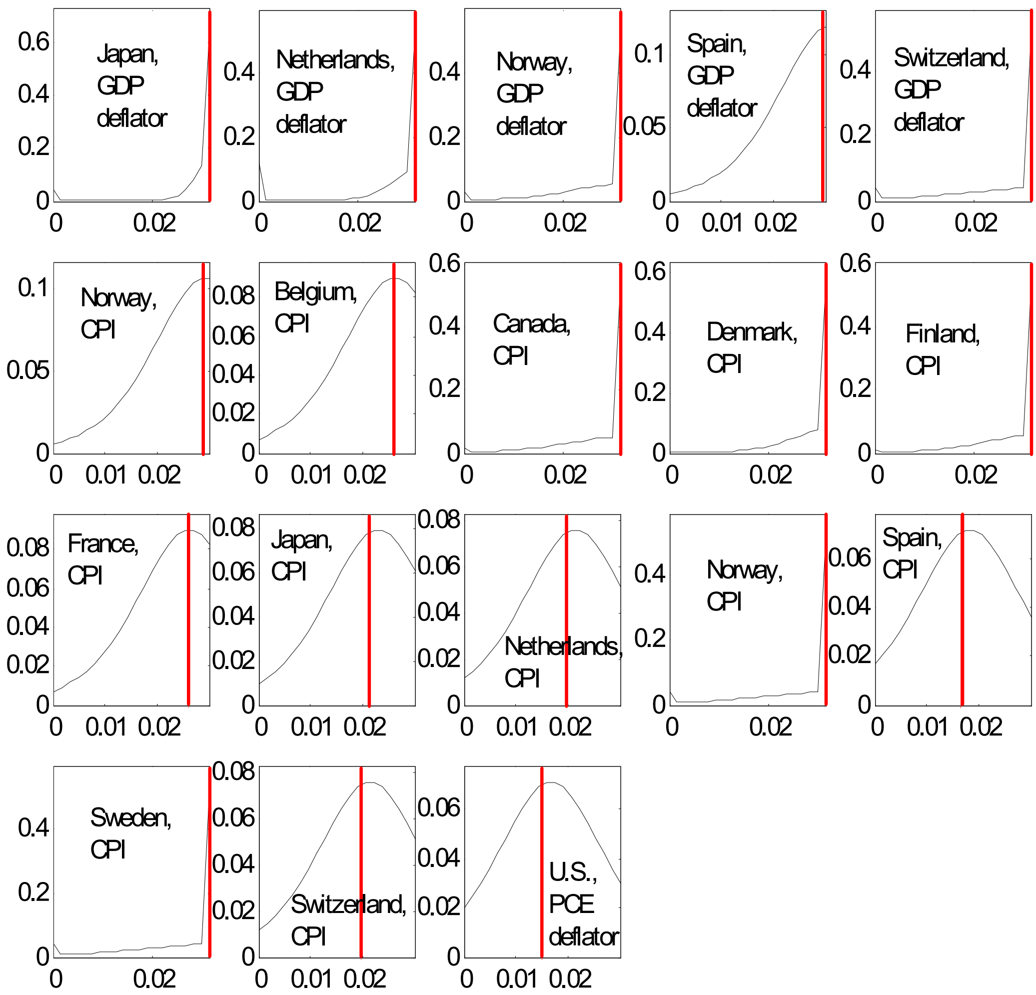

Figure 1: Deconvoluted probability density functions for $\hat{\lambda}$ (the vertical red bar is the median-unbiased estimate of $\hat{\lambda}$; for details on the deconvolution procedure, see text)

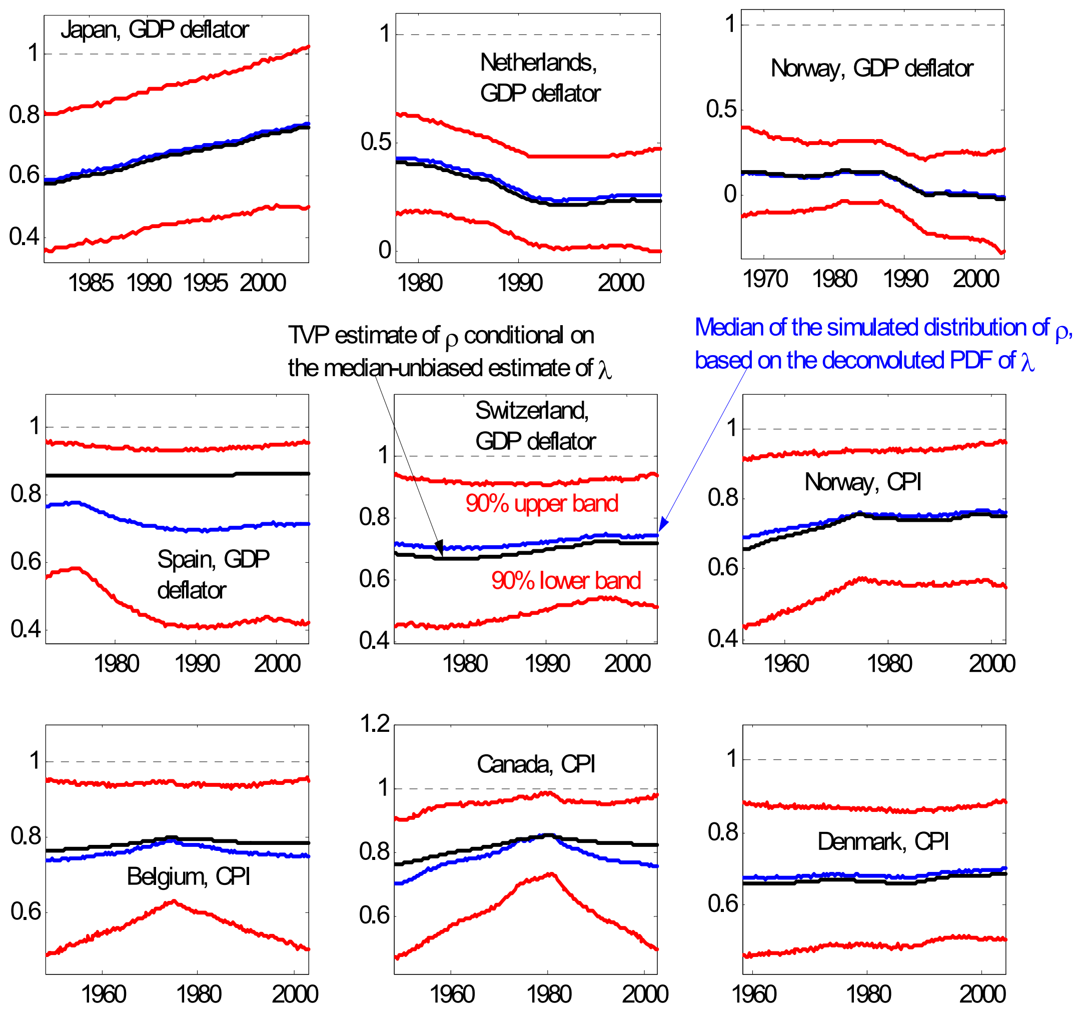

Figure 2: Time-varying parameters estimates of ρ, and 90% confidence bands

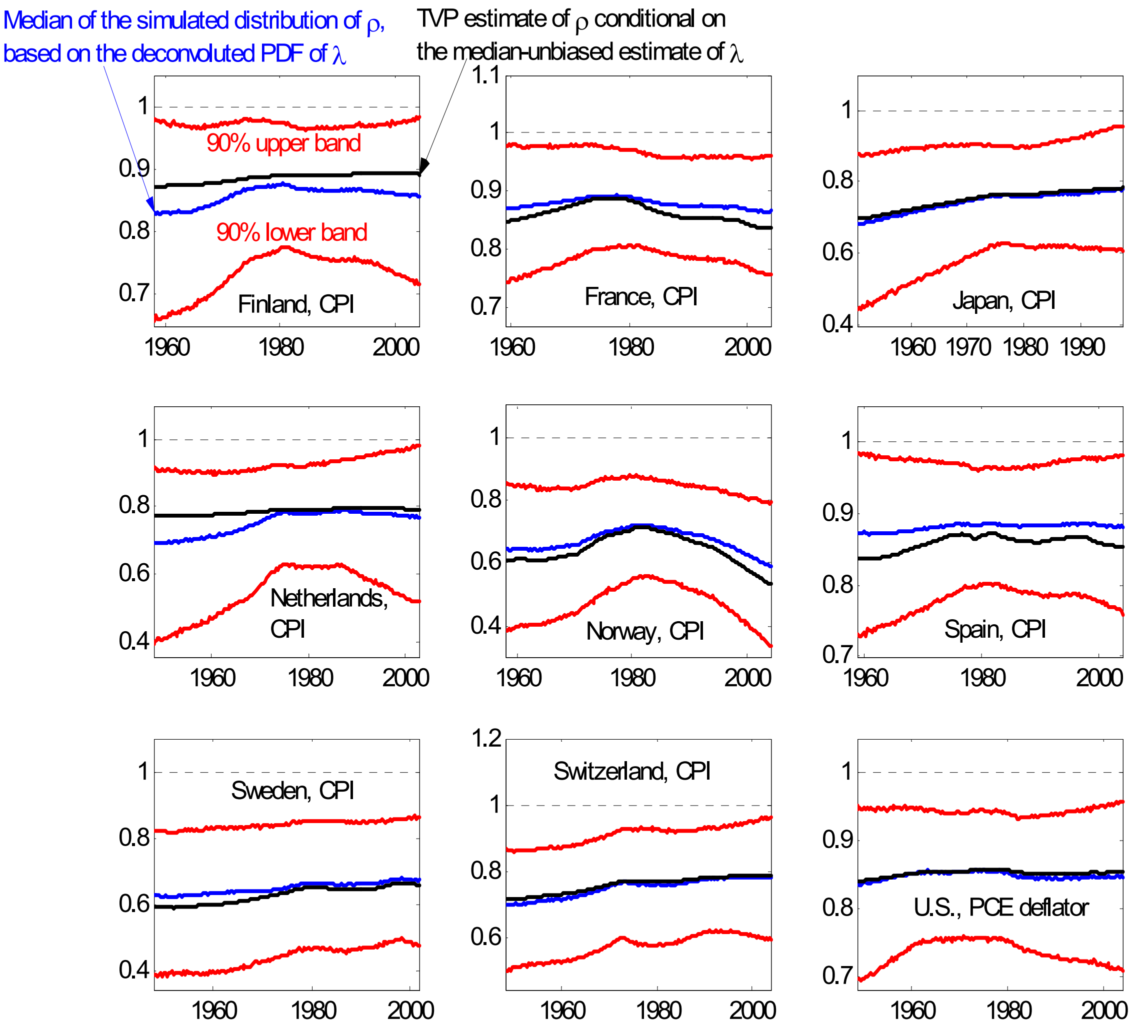

Figure 3: Time-varying parameters estimates of ρ, and 90% confidence bands (continued)

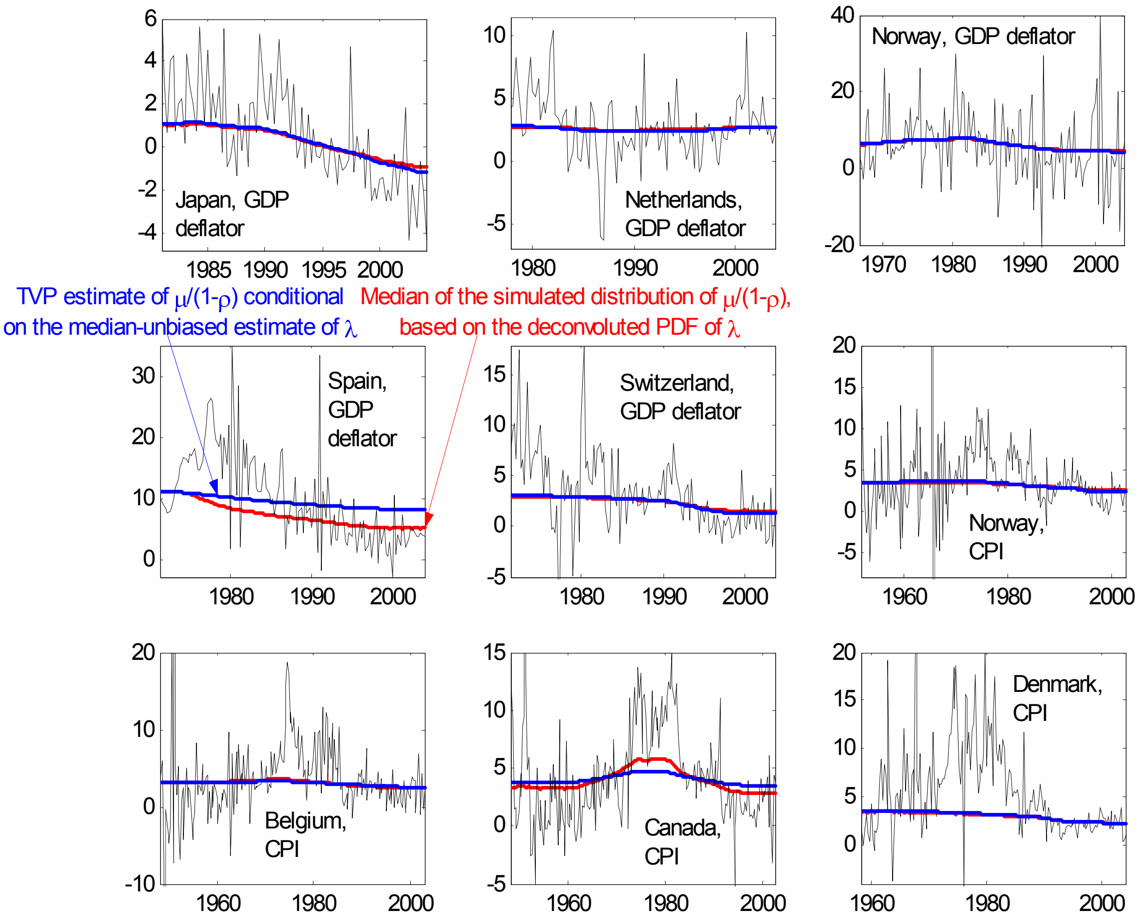

Figure 4: Time-varying parameters estimates of mean inflation

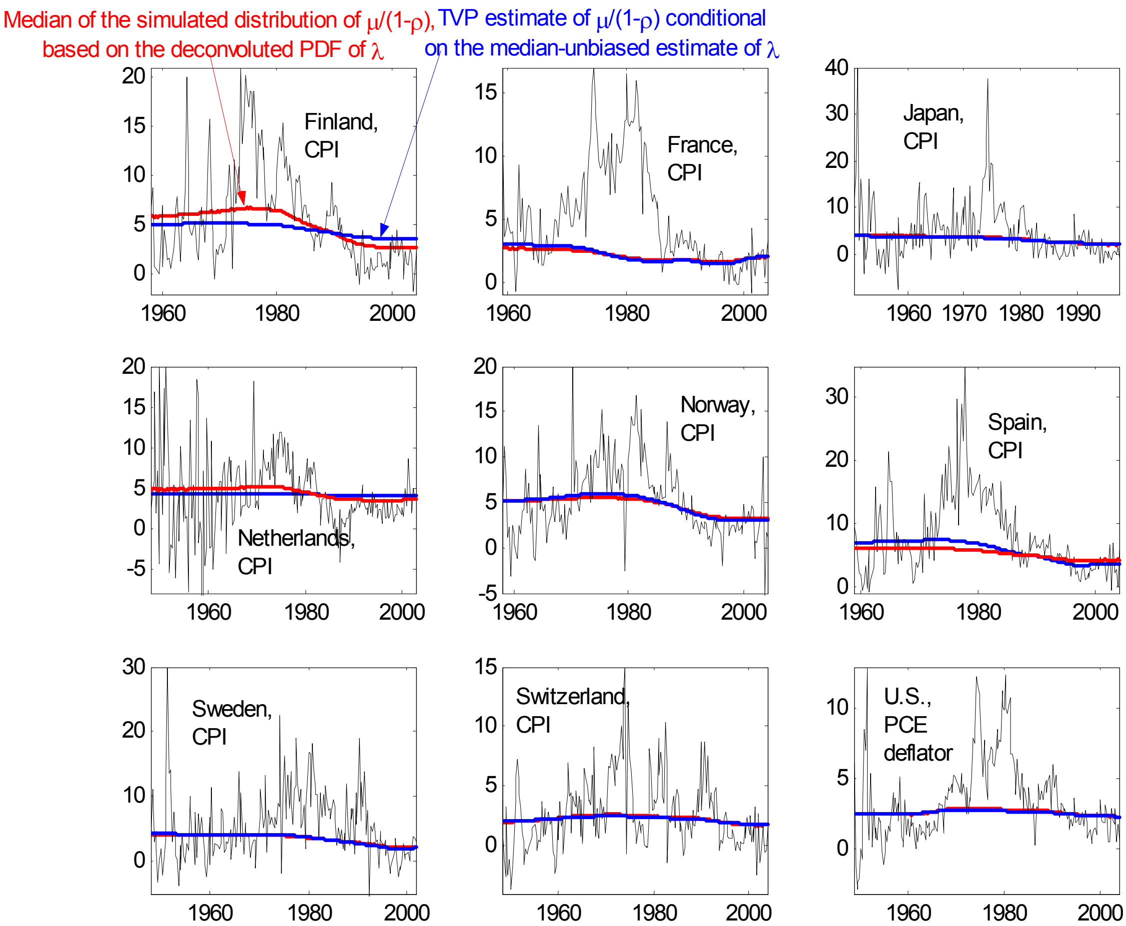

Figure 5: Time-varying parameters estimates of mean inflation (continued)