The Persistence of Inflation in the United States*

Frederic Pivetta and Ricardo Reis†

McKinsey & Company and Princeton University

Abstract

Has the persistence of inflation in the United States changed since 1965? We estimate the persistence of inflation over time using different measures and estimation procedures and we produce confidence intervals for our estimates as well as formal tests of unchanged persistence. We find that inflation persistence has been high and approximately unchanged in the U.S. over our sample period. We reconcile our results with other studies that reached different conclusions.

JEL classification: E31; C22; C11.

Keywords: Inflation persistence; Bayesian time-varying parameters; Median unbiased estimates; Structural stability.

---

*We are grateful to Gary Chamberlain and Jim Stock for discussions, and to Tim Cogley, Peter Ireland, Greg Mankiw, and anonymous referees for comments. Pivetta gratefully acknowledges financial support from the College InterUniversitaire en Management, and Reis from the Fundação Ciência e Tecnologia, Praxis XXI.

†Corresponding author: Ricardo Reis, 324 Bendheim Hall, Princeton University, Princeton, NJ 08544, USA. Tel.: +1-609-258-8531. E-mail address: rreis@princeton.edu.

1 Introduction

While inflation is an important macroeconomic variable that has received much attention, there is still controversy on its properties. This paper is a contribution to the measurement over time of a key property of inflation: its persistence.

We start our investigation by estimating a Bayesian non-linear model of inflation dynamics and using it to provide a time series for inflation persistence in the United States. Our model is similar to that in the innovative contribution by Cogley and Sargent (2002). They find that the persistence of inflation in the United States rose in the early 1970s and remained high during this decade, before starting a gradual decline from the early 1980s until the present (similar to the results of Taylor, 2000, and Brainard and Perry, 2000). We depart from their study in three important ways. First, while they displayed only point estimates, we also present the Bayesian credible sets. Second, whereas they condition their priors and restrict the parameter space solely to stationary representations, we allow for the possibility that inflation may have a unit root, a possibility that the data do not reject. Third, we compute alternative statistical measures of persistence, especially to try to distinguish between changes in volatility and changes in persistence. The first of these modifications is enough to alter the Cogley and Sargent conclusion. All three together lead to a different conclusion: inflation persistence in the United States is best described as unchanged over the last three decades.

We then take a different approach to the problem by estimating a model of inflation dynamics grounded on classical statistical theory. Using local-to-unity asymptotics, we compute median unbiased estimates for different measures of inflation persistence. We find again that inflation persistence is best described as unchanged since 1965. Classical estimates also provide us with formal tests of the null hypothesis of no change in persistence. We are never able to reject this null.

The results in this paper inform many debates in macroeconomics. Closely related is the debate on whether monetary policy could repeat the mistakes of the 1970s, when inflation spiralled up. Hall (1999) and Taylor (1998) noted that if estimates of the persistence of inflation are revised downwards, tests of the natural rate hypothesis in the spirit of Solow (1968) and Tobin (1968) will start rejecting long-run monetary neutrality.[1] If the central bank feels encouraged to exploit an illusory inflation unemployment trade-off, the result could be high inflation without any accompanying output gains. As the model in Sargent (1999) illustrates, if the policymaker gradually learns about the natural rate hypothesis by successively applying over time the Solow-Tobin test on the available data, a fall in the persistence of inflation will lead to a shift to a high-inflation time-consistent equilibrium. Even if the underlying true model exhibits long-run monetary neutrality, the inflation bias equilibrium becomes self-confirming as inflation endogenously lacks persistence and the central bank keeps rejecting the natural rate hypothesis. In this model, inflation persistence is key in determining whether there is high inflation (as in the 1970s) or low inflation (as in the 1990s). Detecting whether persistence has recently fallen is key in assessing the likelihood of recidivism by the central bank.

On a different front, research on dynamic price adjustment has emphasized the need for theories that generate inflation persistence (see Taylor, 1999). This paper evaluates whether the models should generate changes over time in this persistence. While it is true that inflation persistence has changed before – Barsky (1987) finds that inflation was close to white noise in the pre-war period but very persistent since the 1960s – it is unclear whether a new change occurred in the last 35 years.

Yet a third line of work has identified changes in the way monetary policy is conducted in the United States. Clarida, Gali and Gertler (2000) and Boivin and Giannoni (2005) have econometrically identified shifts in the parameters of Taylor rules fit to describe the setting of interest rates by the Federal Reserve. It remains an open question whether there have been corresponding changes in the response of inflation to monetary shocks, including the persistence of this response. Furthermore, recent research has detected a fall in the volatility

[1] It has long been known though (Sargent, 1971) that if agents form expectations rationally, these tests cannot determine whether the natural rate hypothesis holds.

of output after 1984 (McConnell and Perez-Quiros, 2000) as well as a fall in the volatility of many other macroeconomic series in the 1990s (Stock and Watson, 2003; Kim, Nelson and Piger, 2004). This suggests the possibility of a structural change in the economy in the last 20 years, which if it occurred, could have effects on the persistence of inflation.

Has the persistence of inflation changed over time? Will tests of the natural rate hypothesis soon start rejecting, raising the dangers of recidivism? Do we need theories that generate persistence in inflation and if so should this be modelled as varying over time? Have the changes in monetary policy and in the behavior of output come with changes in inflation persistence? To answer these questions, a necessary first step is to measure the persistence of inflation and to assess whether it has changed over time. This is the aim of this paper.

In section 2, we describe the statistical model and the measures of persistence we will employ. In section 3, we present and discuss the data, and in section 4, we implement a Bayesian approach to estimation of persistence. Section 5 instead provides classical median unbiased estimates of persistence and tests for the null of unchanged persistence. Section 6 compares our findings with those of other papers, and section 7 concludes.

2 Statistical model and measures of persistence

We estimate the following k^{th} order autoregressive representation for inflation:

$$\pi_t = \theta_{0,t} + \sum_{i=1}^{k} \theta_{i,t}\pi_{t-i} + \varepsilon_t. \tag{1}$$

It explicitly allows for changes in persistence over time driven by changes in the time-varying $k+1$ parameter vector $\theta_t' = (\theta_{0,t}, \theta_{1,t}, ..., \theta_{k,t})$. We choose to focus on a univariate model, since persistence is itself a univariate property. Our estimates of persistence at time t will reflect what inflation is expected to be at time $t+s$, conditional on all the present and past values of inflation from time t backwards. Including other variables would lead to an assessment of *predictability*—what do we forecast inflation to be at some future date, given the values of a set of forecasting series today and in the past. Our focus on persistence, not predictability, leads us to work with a univariate model.[2]

By persistence, we mean the long-run effect of a shock to inflation—given a shock that raises inflation today by 1%, by how much do we expect it to be higher at some future date and how long (if ever) will it take to return to its previous level. This concept is intimately linked to the impulse response function of inflation, yet the impulse response is not a useful measure of persistence since it is an infinite-length vector. There are many useful scalar functions of the parameter vector θ_t that serve as measures of persistence. We focus on three, all with some virtues and some faults.

First, is the largest autoregressive root (LAR), which we denote by ρ. If L denotes the lag operator $L^j x_t = x_{t-j}$, the lag polynomial associated with equation (1) can be written as $1-\theta_1 L-...-\theta_k L^k = (1-\rho L)(1-b_1 L)...(1-b_{k-1}L)$, where ρ, b_1, ..., b_{k-1} are the autoregressive roots of which ρ is the largest. In the distant future, the impulse response of inflation to a shock is dominated by the largest root, so the size of ρ is a key determinant of how long the effects of the shock will persist. When $\rho = 1$, the process is infinitely persistent, since after a shock, inflation never returns to its initial level (there is a unit root); when $\rho = 0$, shocks die away immediately and inflation is serially uncorrelated. In between 0 and 1, the higher is ρ, then the longer (to a first approximation) it will take for inflation to fall back to its original level after a shock, and thus the larger is the persistence of inflation. The LAR has been used to measure persistence by Stock (1991) and De Jong and Whiteman (1991) in the context of testing for the presence of unit roots. The main problem with this measure, voiced by Phillips (1991) and Andrews (1993), is that it ignores the effects of the other roots. All else equal, an AR(2) process with roots 0.9 and 0.8 is more persistent than an AR(2) with roots 0.9 and 0.1. While the LAR may provide a good approximation to persistence, considering more roots will provide better approximations.

A second measure of persistence is the sum of the coefficients (SUM) in the autoregressive

[2] In the literature, these two different concepts are sometimes referred to as conditional or unconditional persistence, referring to whether we are conditioning on other variables or not.

process, defined as $\gamma = \sum_{i=1}^{k} \theta_i$. The rationale for this measure comes from realizing that for $\gamma \in (-1, 1)$, the cumulative effect of a shock on inflation is given by $1/(1-\gamma)$. A larger γ therefore intuitively corresponds to higher persistence of inflation. This measure has a history in macroeconomics since the Solow-Tobin test of the natural rate hypothesis is a test for the null hypothesis that $\gamma = 1$, so that low values of γ would provide support to the possibility of recidivism pointed in the introduction.[3] The main problem with this measure is that it is large for a process with an impulse response function where inflation rises quickly to large levels to fall steeply back to zero, than for a process with a slowly decaying impulse response but which increases by little in the beginning, in spite of the second being intuitively more persistent.

Our third measure is the half-life (HL), defined as the number of periods in which inflation remains above 0.5 following a unit shock. We denote it by η, defined by $\eta = \{t \in [0, 40] : E[\pi_t - \pi_{t-1} \mid \varepsilon_0 = 1] \geq 0.5,\ E[\pi_{t+1} - \pi_t \mid \varepsilon_0 = 1] < 0.5\}$, and $\eta = 40$ if the previous set is empty. We interpret half-lives of 40 as infinite persistence since in practice in our application, half-lives above 40 usually indicate non-stationarity and thus an half-life of infinity. This measure of persistence is especially popular in the vast literature that measures deviations from PPP surveyed in Rogoff (1996). Its virtues are that it is intuitive, simple, and easy to associate with the concept of persistence discussed above. There are many arguments against it though. First, if the impulse response function is oscillating, the half-life will severely understate the persistence of the process. Second, even if the impulse response monotonically decays, the half life will be lower if the impulse response is more pronouncedly convex to the origin, when this convexity has little relation to persistence. Third, for very persistent processes, the half-lives are very close to infinity making it difficult to distinguish changes in persistence over time. Finally, there has been less attention in econometric theory devoted to studying the distribution of estimates of the half-life, as opposed to the other two measures of persistence.

[3]Note that the Solow-Tobin test includes unemployment in the estimated regression, so it is not quite the same as this test.

Finally, we must choose the order of the autoregression. There is no clear statistical criterion to choose k, since we will use different estimation methods on different samples. We set $k = 3$ based on the Bayesian Information Criterion (BIC) on the full sample and several different sub-samples.[4]

3 Data description

We use seasonally adjusted quarterly data on the GDP deflator from the second quarter of 1947 to the third quarter of 2001, obtained from the Bureau of Economic Analysis, as our measure of the price level (P_t).[5] Inflation is the quarterly change of the price level at an annualized rate calculated as $\pi_t = 400 \ln(P_t/P_{t-1})$. The plot of the data is in figure 1, which shows the well-known trends. Starting from low levels in the 1950s and 1960s, inflation rose throughout the 1970s, peaking at 1975 at 11.8%. It then remained high, peaking again at 10.5% at the end of 1980. Paul Volcker, then chairman of the Federal Reserve, embarked in an aggressive policy to reduce inflation. As a result, inflation fell to 3% by the beginning of 1983, and feel further to 1.6% by 1986. Throughout the 1990s, it remained stable between 1% and 3%. Table 1 reports the sample means and the standard deviations per decade. The 1970s were a decade of unusually high inflation while, after an intermediate decade in the 1980s, the mean inflation in the 1990s was back to the values observed in the 1950s and 1960s. The 1990s differed from previous decades in that the volatility of inflation was lower than ever. The 1970s on the other hand are not unusual in terms of the standard deviation of inflation.

[...FIGURE 1 ABOUT HERE...]

[...TABLE 1 ABOUT HERE...]

[4] We also computed most of the estimates with $k = 5$ and found no significant changes.

[5] We also conducted most of the analysis using the CPI. There were no substantial changes to the conclusions.

A natural first look at the persistence of inflation is to see the evolution over time of the sample first-order serial correlation. Table 1 shows the serial correlation per decade. We see that in the 1950s, it was unusually low, mostly due to the Korean war price controls at the beginning of the decade. After rising from the 1970s to the 1980s, in the 1990s the serial correlation of inflation was down to the 1970s value. Figure 2a shows rolling estimates of the serial correlation, for which the estimate at t is based on data from $t-55$ to t, a window of 14 years.[6] The same broad trends are again visible. For most of the 1960s, the serial correlation was quite low, only rising right at the end to the 0.81 value reported in Table 1. The serial correlation during the 1970s was around 0.7, while in the 1980s it rose to peak at 0.93. Then, after another sharp fall down to 0.4 in 1995, the serial correlation rose until 1998 up to 0.7, since when it has been falling again. Figure 2b instead shows recursive estimates. The estimate at t now incorporates all the information from the beginning of the sample up to time t. A different picture emerges: After starting at around 0.4 and staying there during the 1960s, the serial correlation shot up throughout the 1970s and the first half of the 1980s, stabilizing around 0.74 with no noticeable change throughout the 1990s.

[...FIGURES 2a AND 2b ABOUT HERE...]

4 Bayesian time-varying parameter estimates

A Bayesian approach is well-suited to this problem. It explicitly treats the parameter vector θ_t as being random and time-varying and provides posterior densities for θ_t at all points in time. From these, one can obtain posterior densities for the measures of persistence.

4.1 The model

Following Cogley and Sargent (2002), we study a state-space model given by:

[6] The choice of 14 years, which will also be followed in other sections of the paper, was guided solely by the need to have long enough samples to assess persistence. We have replicated most of the results with 12 and 10 year samples, with no noticeable changes.

(1) the measurement equation in equation (1);

(2) an equation for the evolution of the state:

$$p(\theta_{t+1} \mid \theta_t, V) \propto I(\theta_{t+1}) f(\theta_{t+1} \mid \theta_t, V), \tag{2}$$

where $p(\theta_{t+1} \mid \theta_t, V)$ is the density of next period's state (θ_{t+1}) conditional on this period's state θ_t and on the covariance matrix of the disturbances to the system V. This density is proportional to the product of a normal density $f(\theta_{t+1} \mid \theta_t, V)$ with mean θ_t and variance V, and an indicator function $I(\theta_{t+1})$, which imposes prior subjective restrictions on the model. These restrictions may lead to non-linearities. However, the posterior distributions are still well-defined and Monte Carlo methods can be used to approximate them. We will use the notation $p(.)$ to designate an arbitrary probability density function and $f(.)$ to refer more specifically to a normal density function. The errors of the measurement and state equations $(\varepsilon_t, v'_{t+1})'$ are i.i.d. normal random variables with mean 0 and covariance matrix

$$V = \begin{pmatrix} \sigma_\varepsilon^2 & C' \\ C & Q \end{pmatrix} \tag{3}$$

of dimension $(k+2)$x$(k+2)$, where σ_ε^2 is the scalar variance of measurement innovations, Q is the covariance matrix of state innovations which is $(k+1)$x$(k+1)$, and C is a cross-covariance vector, $(k+1)$x1. In maybe more familiar form, the state equations are given by $\theta_{t+1} = \theta_t + v_{t+1}$ with the added truncation $I(.)$ to possibly restrict the parameter vector to some class. In empirical Bayes terminology, θ_t are the parameters and the elements of V are the hyperparameters.

The priors for θ_0 are independent from the hyperparameters and also follow a (possibly truncated) Gaussian distribution, with mean $\bar{\theta}$ and variance $\bar{P}$. The hyperparameters come from an inverse-Wishart distribution with scale matrix $\bar{V}^{-1}$ and T_0 degrees of freedom.[7] The

[7] The Wishart distribution is the multivariate analogue of the chi-squared distribution.

joint prior distribution is hence given by:[8]

$$p(\theta_0, V) \propto I(\theta) f(\bar{\theta}, \bar{P}) IW(\bar{V}^{-1}, T_0). \tag{4}$$

The values of $(\bar{\theta}, \bar{P}, \bar{V}, T_0)$ come from estimating a time-invariant autoregression on an initial sub-sample 1948.1-1958.4. To estimate the model, we use data from 1959.1 to 2001.3, starting the estimates at 1965.1. We denote the partial histories of the variables as $\Pi^T = [\pi_1 ... \pi_T]'$ and $\theta^T = [\theta_1 ... \theta_T]'$ and the potential futures from date T to $T + H$ as $\Pi^{T+1,T+H} = [\pi_{T+1} ... \pi_{T+H}]'$ and $\theta^{T+1,T+H} = [\theta_{T+1} ... \theta_{T+H}]'$.

Our aim is to obtain a posterior predictive density for the parameters. We can decompose the joint posterior into two fundamental components, using Bayes law:

$$\begin{aligned} & p(\Pi^{T+1,T+H}, \theta^{T+1,T+H}, \theta^T, V \quad | \quad \Pi^T) = \\ & p\left(\theta^T, V \mid \Pi^T\right) p(\Pi^{T+1,T+H}, \theta^{T+1.T+H} \quad | \quad \Pi^T, \theta^T, V). \end{aligned} \tag{5}$$

The first term captures beliefs about the past—it is the joint posterior density for both the past states and the hyperparameters. The second term involves beliefs about the future, conditional on the beliefs about the past. This includes beliefs not only about the future data observations but also about the future evolution of the state vector. This suggests a two-step procedure to obtain future draws that we can use to approximate the joint posterior distribution. First, obtain by multiple draws the posterior of past states and hyperparameters. Then, conditional on each past draw, draw future states and inflation paths. This gives draws from this posterior density. The first part of the algorithm uses the Gibbs sampler to simulate draw of $\left(\theta^T, V\right)$ from $p\left(\theta^T, V \mid Y^T\right)$. The second part runs the system forward successively drawing state and measurement innovations from their known distributions. The Appendix describes the exact steps.

[8] This choice of priors coincides with that of Cogley and Sargent (2002) to allow comparison of our results. This is not an innocuous choice though. See Uhlig (1994) for the implications of this prior and possible alternatives.

Implementation involves choosing the number of iterations to use in our approximation of the posterior density. We draw 420 past histories, discarding the first 120 to remove dependence from the starting point of the Gibbs sampler, and then draw 400 futures per history. We set $H = 120$ quarters so that we draw futures with 30 years duration. Thus, at each date we have in total $400 * (420 - 120) = 120,000$ future trajectories for inflation (π_t) and the states (θ_t), each 120 quarters long. Because this focuses on future draws, the estimates of persistence are forward-looking, since they capture expectations of the future persistence of inflation as of a point in time. For each trajectory of future data, we compute our measures of persistence ρ, γ and η, using an AR(3) as a base. We then take the median over the trajectories and the 5% and 95% percentiles to give the Bayesian 90% credible set.[9]

4.2 Replicating Cogley and Sargent (2002)

Our specification differs from Cogley and Sargent's (2002) in many ways. First, they estimate a VAR in order to address other issues aside from inflation persistence.[10] Second, they argue that non-stationary representations for inflation are implausible from the perspective of economic theory, since they imply an infinite asymptotic variance for inflation, which could never be optimal if the central bank's loss function includes the variance of inflation. Accordingly, they truncate the parameter space using the indicator function $I(.)$ to exclude non-stationary representations for inflation. Third, they choose to measure persistence by the spectrum at frequency zero, which we denote by S_0. For each trajectory of future inflation, they estimate a spectrum and take the median of these estimates at frequency zero.

Figure 3a shows the median estimates and Bayesian credible sets of S_0. To ease interpretation, figure 3b shows the same estimates using a logarithmic scale in the vertical axis. Our

[9] There are severe computational constraints on this exercise. Simultaneously using 5 computers with approximately 1700 Mhz, it takes about one month to run the full agorithm. These constraints explain our choice to settle for a number of draws that is relatively small for what is standard in Gibbs sampling. Section 4.5. examines the robustness of the results to the number of draws.

[10] Cogley and Sargent (2002) are also concerned with making inferences on long-run forecasts of inflation, tests of the natural rate hypothesis, and estimates of Taylor-type policy rules. To do so, they estimate a VAR on three variables: inflation, unemployment and interest rates. The focus of this paper on inflation persistence leads us to work instead with an AR.

estimates are similar to those in Cogley and Sargent (2002). The estimates of S_0 rise until the 1970s, remaining at high levels during this decade, and start a gradual decline since 1981 until the present. Unlike them, we nevertheless also plot the credible sets. These turn out to be quite large, so that paths with constant inflation persistence are perfectly consistent with the data. Thus, once we take parameter uncertainty into account, there is no longer decisive evidence that inflation persistence has changed in the last three decades.

[...FIGURES 3a, AND 3b ABOUT HERE...]

4.3 Forward-looking measures

The use of S_0 in this application is nevertheless problematic, since it is likely to confuse changes in persistence with changes in volatility. From table 1, we have strong signs that the variance of inflation has been falling in the last 20 years. A fall in the variance would lower the spectrum of inflation at all frequencies, including 0. Thus, S_0 would fall even if all the other autocovariances, and so persistence, were unchanged. To see this, note that if the model consisted solely of the measurement equation, then:

$$S_0 = \frac{1}{2\pi}\frac{\sigma_\varepsilon^2}{(1-\gamma)^2}. \tag{6}$$

Even if a variance-independent measure of inflation persistence as the sum of the coefficients on inflation γ is unchanged, S_0 will fall with the variance of inflation. Indeed, Cogley and Sargent (2002) note that their estimates of S_0 are highly correlated with their measures of the variance of inflation.[11]

Figures 4a, 4b, and 4c depict the median estimates and 90% credible sets for the largest autoregressive root (ρ), the sum of lagged coefficients on inflation (γ), and the half-life (η).

[11] An alternative measure of persistence, not subject to this criticism, would be the normalized spectral density at frequency zero. The ideal though would be to have a statistical model that allowed for time-varying volatility. Our model in section 4.1 does not take this approach, because of the severe computational constraints that we faced. We hope that future research pursues this.

The three measures give the same message: inflation persistence is basically unchanged over the last 30 years, with at most a very slight fall in the median estimates after 1980. The credible sets cover a large area and constant paths for inflation persistence are consistent with the data over a wide range. Compared with figures 4a and 4b, figure 4c shows smoother estimates, due to the non-linearity of the half-life estimator.

[...FIGURES 4a, 4b, AND 4c ABOUT HERE...]

Another issue in the Cogley and Sargent (2002) application is the truncation of the parameter space to exclude non-stationary representations. If there is a strong belief that inflation is stationary, this is the appropriate way to introduce these strong subjective priors of the researcher into the model. However, it comes with some potentially undesirable implications. For the sake of illustration, say that $k = 1$, so that we are working with an AR(1) and the parameter θ_1 equals both ρ and γ. If at a point in time our estimate of θ_1 is 0.99 with standard deviation 0.05, the restriction to stationary draws next period implies that approximately 42% of the draws are discarded. The implied distribution for the parameter then becomes similar to the lower half-tail of the normal, scaled up to have measure one. It is very asymmetric (we are putting only a 14% chance on the parameter rising and 86% on it falling), and this could be seen in figure 4 where the median is much closer to the upper bound of the credible set than to the lower bound. Moreover, by placing such a strong weight on declines, θ_{t+1} has a distribution that does not have a mean of $\theta_t = 0.99$, but instead a mean of 0.956. Thus, we should expect that the truncation strongly pushes the results towards low values of persistence.[12]

Furthermore, the argument that we should exclude the possibility of an infinite asymptotic variance of inflation is not entirely convincing. If the objective function of the policymaker is the expected discounted sum of the variance of inflation, then as long as the

[12] Sims (2002) discusses the economic implications of this restriction, namely in biasing tests of the natural rate hypothesis ($\gamma = 1$) towards rejection.

discount factor is smaller than one, even if inflation has a unit root, the loss is finite. Moreover, according to this argument, if output also enters the central bank's loss function we should also discard explosive representations of output, ruling out by assumption unit roots in output despite the large literature arguing for their presence. If it is the output gap that enters the central bank's objective function, the Cogley and Sargent argument would allow for unit roots in output, but it is unclear why there should not also be an inflation gap in the objective function rather than actual inflation, in which case unit roots in inflation should not be excluded. More generally, taking the Cogley and Sargent restriction to its logical extreme, any variable that enters the objective function of risk-averse agents must be stationary. This includes most variables of interest in economics.

We proceed by re-estimating the model without imposing the stationarity restriction. This amounts to setting $I(\theta_{t+1}) = 1$ always. Estimates and 90% credible sets for our three measures of persistence are in figures 5a, 5b and 5c.[13] Note that the estimates are substantially larger than before, and we generally cannot exclude the possibility of explosive representations. Concerning changes, our earlier result remains: there is no detectable change in the median estimates of inflation persistence in the last 30 years, and a wide range of values of persistence lie at all points in time within the 90% credible set.

[...FIGURES 5a, 5b, AND 5c ABOUT HERE...]

In taking as the estimator of persistence the median of the posterior distribution, we are implicitly choosing an absolute error loss function. A common alternative is a squared error loss, in which case the Bayes estimator is the mean of the posterior density. Figures 5a, 5b and 5c also plot the mean. The conclusion of no evidence for a change in persistence remains, as the mean closely tracks the median and is stable over time and across measures of persistence.

[13] Using S_0 to measure persistence would only be valid if the process for inflation was covariance stationary, but this is not the case for our three preferred measures of persistence.

Finally, note from figure 5b that since 1980 there is a progressive fall in the lower bound of the credible set. A little of this fall (though much less so) also happens in the posterior for the largest autoregressive root in figure 5a. We inspected the evolution of the posterior distributions over time and found that, below the 40% percentile, the distributions have become somewhat flatter in the last 20 years. This could be interpreted as some evidence that persistence has fallen as the probability of lower values has increased. Yet this is at best weak support for the hypothesis of a change in persistence. Since the remaining pieces of evidence go against the hypothesis, we conclude persistence is best described as unchanged over time.

4.4 Backward-looking measures

The forward-looking measure of persistence we have employed so far are meant to capture the perspective of a policymaker who at a point in time is trying to foresee what the persistence of inflation in the future will be. An alternative is to estimate our Bayesian time-varying parameters model for the whole sample, and use the posterior distribution of the models' parameters at each point in the sample to generate posterior distributions of the inflation persistence measures at each point in time. These are backward-looking measures of persistence, since they are the estimates of persistence that the econometrician forms at a point in time, given all of the sample data until then.

Figures 6a to 6c present backward-looking estimates of persistence. Because these are less computationally intensive, they use a much larger number of draws: 30,000, discarding the first 10,000. These backward-looking measures are less decisive on the question of whether persistence has changed.[14] On the one hand, there is a sharp fall in the median estimates of persistence in 1975 and 1982. Moreover, there are clear trends in the median estimates of persistence, rising from 1960 to 1975, and slowly declining since 1990. On the other

[14] The half-life estimates are tightly concentrated between 2 and 3 quarters, and almost always unchanged despite the large variations in the LAR and the sum of autoregressive coefficients. This is likely driven by the non-linearity of half-life estimates and serves as an illustration of the possible danger of focusing solely on this measure.

hand, the credible sets are wide and consistent with the view that persistence if unchanged. Furthermore, the rise from 1960 to 1975 can be explained by the choice of priors, which have very low persistence and thus exert a strong influence in the estimates in the beginning of the sample. In the next section, we show that having priors with higher persistence eliminates this rise. As for the sharp falls, the first is clearly associated with the oil price shock and the second with the Volcker disinflation. Noticeably though, after these two isolated episodes, the estimates return to their previous levels.

[...FIGURES 6a, 6b, AND 6c ABOUT HERE...]

The measures of persistence since 1990 leave open the question of whether persistence has recently fallen. As with the forward-looking measures, there is an increase in the mass of the distribution at lower values of persistence. The credible sets widen to now include many values of persistence that have quite different implications. However, it is unlikely that a researcher coming into the 1990s with priors mildly in favor of unchanged persistence would change her views. The case for a change in persistence is slim.

4.5 The role of the priors and the number of draws

The priors for our application were obtained from a time-invariant autoregression on an initial sub-sample of the data from 1948.1-1958.4. Yet, as we observed in section 2, this period is somewhat unusual in that the serial correlation of inflation was much lower than what it typically has been since 1960. The LAR for the prior parameters is 0.54 lying with 90% prior probability on the (0.39, 0.80) interval, well below most of the estimates in figures 4 to 6.

To check on the robustness of our results to different priors, we undertook two experiments. First we changed the priors for $\bar{\theta}$ to raise the LAR while keeping the priors for $(\bar{P}, \bar{V}, T_0)$ as before. The new implied LAR was 0.95, more in accord with the post-war experience, lying with 90% probability in (0.82, 1.05). Our second experiment consisted of using as priors the estimates from a time-invariant regression on the data from 1959.1 to 2001.3. These are, of course, not valid priors since they use future information. The thought experiment we had in mind was to consider the priors of a researcher that comes into this exercise with her experience of studying inflation in the last 40 years. This researcher has much tighter priors, with an LAR that lies, with 90% probability, in the interval (0.91, 0.98).

Figures 7a to 7c show the estimated backward-looking measures of persistence, the left panel corresponding to the first experiment, and the right to the second experiment. The contrast between figures 7 and 6 is remarkable. The estimates of persistence are now typically unchanged over the entire sample period. With the very uninformative priors, the posterior credible set is wide, and the mass in the left tail increases in the 1990s as before. With the second set of priors, the credible sets are remarkably tight throughout with very little variation over time.

We were not able to repeat this experiment with the forward-looking measures because of our computational constraints. Still, given the results with the backward-looking measures, it is reasonable to expect that using priors with very little persistence tends to influence the results towards finding changes in persistence. Since with these priors, we still found that persistence is best described as unchanged, this suggests that using priors with higher persistence would reinforce this conclusion.

[...FIGURES 7a, 7b, AND 7c ABOUT HERE...]

Another potential concern in the implementation is that we used a small number of draws in the forward-looking estimates. Even though this issue is worrisome, taking more draws was infeasible given our computational constraints. One way we checked whether this is problem was by estimating the backward-looking measures, which are much less computationally demanding, for different number of draws. Figure 8 shows estimates of the LAR. Panel a) uses 420 draws discarding the first 120, as in the forward-looking measures. Panel b) replicates figure 6a, which used 30,000 draws, discarding the first 10,000. Panels c) and d) use an intermediate number of draws: 11,000 discarding the first 1,000, and 10,000 discarding the first 5,000 respectively. The figure shows that using a limited number of draws leads to estimates that are similar to those with many more draws. While this is only suggestive, it indicates that the small number of draws behind the forward-looking estimates in figures 4 and 5 may have been enough.

[...FIGURE 8 ABOUT HERE...]

5 Classical median unbiased estimates

Classical econometrics provides an alternative set of estimation techniques for persistence in our model. We pursue these in this section. Now, the parameter vector is treated as fixed, but variability in the sample data is taken into account. Consequently, the estimates that follow are different in spirit from the Bayesian estimates of the previous section, and could easily lead to different conclusions, as it happens in the related literature on unit roots.[15]

5.1 Estimates

Alternative estimators of persistence and associated asymptotic distributions are provided by Stock (1991) and extended by Andrews (1993) and Andrews and Chen (1994). They focus on the class of median unbiased estimators, defined as the estimators $\hat{\theta}$ of the true parameter θ, such that the median of $\hat{\theta}$ equals θ for all possible θ in the parameter space. Median unbiased estimates are generally well above least squares estimates, since the latter are known to be biased downward when the parameters are close to a unit root.

The approach consists of re-estimating equation (1) on different sub-samples of the data, taking care to obtain median unbiased estimates of persistence for each regression. Stock (1991) used local-to-unity asymptotics to establish the asymptotic confidence interval for the

[15] See the 1991 issue of the Journal of Applied Econometrics for a contrast between Bayesian and classical approaches in testing for unit roots.

LAR. For a fixed sample in which the parameter vector is treated as constant, equation (1) becomes:

$$\theta(L)\pi_t = \theta_0 + \varepsilon_t, \tag{7}$$

where $\theta(L)$ is a k order lag polynomial, $\theta(L) = 1 - \sum_{j=1}^{k} \theta_j L^j$, that can be decomposed into $\theta(L) = (1-\rho L)b(L)$, and ε_t is a martingale difference sequence, homoskedastic ($E\left[\varepsilon_t^2\right] = \sigma^2$), and with finite fourth moments ($sup_t E\left[\varepsilon_t^4\right] < \infty$). The local-to-unity model for the LAR is $\rho = 1 + c/T$, for a fixed scalar c, so that ρ is within a $1/T$ neighborhood of 1, where T is the sample size. The polynomial $b(L)$ is a lag polynomial of order $k-1$ with all stable roots.

Rearranging, we obtain the Augmented Dickey-Fuller regression:

$$\pi_t = \theta_0 + \gamma\pi_{t-1} + \sum_{j=1}^{k-1} \phi_{j-1} \triangle\pi_{t-j} + \varepsilon_t. \tag{8}$$

The parameters ϕ_j are given by $\phi_j = -\sum_{i=j+1}^{k} \alpha_i$ and the α_i are the components of the $k-1$ order lag polynomial defined by $\alpha(L) = L^{-1}(1-(1-\theta(L))$. The coefficient on lagged inflation is our familiar measure of persistence, the sum of the autoregressive coefficients γ, which in this parametrization equals $\gamma = 1 + cb(1)/T$. Estimating this regression by least squares, we can compute the t-statistic for the hypothesis that $\gamma = 1$, commonly known as the Augmented Dickey-Fuller (1979) statistic. Using the tables in Stock (1991) and following the procedure in Kendall and Stuart (1967), we obtain a median unbiased estimate and a confidence interval for c, from where a confidence interval for ρ follows from $\rho = 1 + c/T$. To obtain an estimate and a confidence interval for the sum of the autoregressive coefficients, γ, we need a consistent estimator of $b(1) = 1 - \sum_{j=1}^{k-1} \phi_{j-1}$. Andrews and Chen (1994) explain how to obtain this iteratively. Starting with the initial least squares estimate of $b(1)$, we compute an initial estimate of γ, denoted $\hat{\gamma}^{(1)}$. Regressing $\pi_t - \hat{\gamma}^{(1)}\pi_{t-1}$ on a constant and the lags in the change in inflation, we obtain a second estimate of $b(1)$ from where a second estimate of γ gives $\hat{\gamma}^{(2)}$. Iterating until convergence, we obtain a median unbiased estimate of the sum of the autoregressive coefficients. As for the half-life, note that as T tends to infinity so does the half-life. Following Rossi (2005), we make the additional assumption that the half-life as a fraction of the sample size (η/T) converges to a finite value as T goes to infinity, in which case we can obtain median unbiased estimates and confidence intervals for the half-life as well.

Median unbiased estimates and confidence intervals for ρ and γ on recursively increasing samples are presented in figures 9a and 9b from 1970 onwards, where the first estimate uses 14 years of data. Also plotted are the median unbiased estimates on the full sample. Again, it is striking how stable and unchanged over time these estimates are. With a single exception, for all points in time, a wide range of values of persistence are consistent with the possibility that persistence has been unchanged over the sample period. The exception is the short period 1974-75, when the confidence intervals tighten considerably. However, in this period, whereas the largest auto-regressive root increases, the sum of autoregressive coefficients falls, so there is no clear sign of a change in persistence.[16]

[...FIGURES 9a AND 9b ABOUT HERE...]

Estimates over rolling 14-year windows are presented in figures 10a and 10b.[17] The short sample over which each regression is estimated leads to more imprecision and time-variation. Still, the broad trends are the same as those obtained on recursive sample, with again unchanged persistence. The exception is an anomalous episode in 1996, with a sharp fall in the estimates.[18] When we re-estimate the model using a 12-year window, the large drop is now in 1994, and with a 10-year window, it occurs in 1992. This points to the source of the anomaly lying in the beginning of the rolling sample in the 1981-1983 period.

[16] We do not report the estimates for the half-lives because the results are easily describe din one sentence. Since the LAR is consistently above one, the median unbiased estimates of the half-life are almost always at the 40 (infinity) upper bound.

[17] We replicated all of the results with 10 and 12 year windows and obtained qualitatively the same results. The only difference is that confidence intervals are even wider due to the shorter samples, which further reinforces our conclusions.

[18] The drop is sharper if one uses the CPI instead of the GDP deflator. In both cases, the median estimate and 10% lower bound are discontinuous because we are outside the range of critical values tabulated by Stock (1991).

This was indeed an anomalous period in the data for inflation, commonly referred to as the Volcker disinflation. We see this episode mainly as an illustration of the danger of using short samples to infer the persistence of a series, especially when history warns us that an extraordinary event took place.

[...FIGURES 10a, AND 10b ABOUT HERE...]

The estimates and confidence intervals above used asymptotic distributions which may be misleading with finite samples. Andrews and Chen (1994) describe a procedure to obtain finite-sample approximately median unbiased estimates of the parameters in equation (8) assuming a normal distribution for the errors ε_t. The procedure is again iterative: given the initial least squares estimates of $\phi_1, ...\phi_{k-1}$, we define the approximately unbiased estimator $\gamma_{AMU}^{(1)}$ to be the value of γ which leads the least squares estimator to have as its median the initial least squares estimate of γ. This involves finding $\gamma_{AMU}^{(1)}$ such that generating data with this value for γ and the estimates of $\phi_1, ...\phi_{k-1}$ for the remaining parameters, and estimating least squares regressions on many sequences of simulated data, the median of these estimates is exactly the value of γ obtained from the initial least squares regression. Conditional on $\gamma_{AMU}^{(1)}$, we obtain new estimates of $\phi_1, ...\phi_{k-1}$, and calculate a new median unbiased estimator $\gamma_{AMU}^{(2)}$. This is repeated until convergence. By an analogous procedure, we can obtain the 5% and 95% bounds for the largest autoregressive root, the sum of autoregressive coefficients, and the half-life. Note that a limitation of this approach is that it constrains γ to lie in the interval $(-1, 1)$. The results for recursive samples are in figure 11a, which is broadly similar to figure 9b. The only noticeably difference is that the estimates are smaller, which is probably driven by the restriction to the interval $(-1, 1)$ imposed by this procedure. Rolling sample results are in figure 11b, which again are similar to figure 10b, with the exception of the 1975 to 1985 period where, contrary to any of the results so far, the path of inflation persistence over time has a V-shape. Still, the estimates are much more volatile and the confidence intervals are much wider. Overall, these approximate estimates suggest there is not much distortion from using asymptotic distributions in our applications.[19]

[...FIGURES 11a AND 11b ABOUT HERE...]

5.2 Formal tests

We employ two tests to inform us on the validity of the null hypothesis of constant persistence.

Banerjee et al. (1992) develop the asymptotic theory for rolling and recursive tests of unit roots. Their null hypothesis is that $\gamma = 1$ for the entire sample period (and so $c = 0$ and $\rho = 1$). This hypothesis is stricter than our null hypothesis of constant inflation persistence which requires that ρ and γ are constant but not necessarily one. The test statistics are the maximum and the minimum of the Dickey-Fuller statistic over the different regressions, $t^{\max}$ and $t^{\min}$, and Banerjee et al. (1992) provide critical values for sample sizes of 100.[20]

For the recursive estimates, the $t^{\max}$ is 1.07 and $t^{\min}$ is -2.74. The respective 10% significance level critical values are -1.73 and -4.00, so in both cases we do not reject the null hypothesis and infer there is no evidence of a change in inflation persistence. For the rolling Dickey-Fuller statistics, the maximal value is 1.68, to be compared with the 10% significance level critical value of -1.73. We therefore do not reject the null of a constant unit root. The minimal statistic is -5.30, which compared with a 2.5% critical value critical value of -5.29, leads to a rejection of the null. However, inspection of the sequence of the statistics reveals that the rejection is entirely driven by the 1996 estimates, which as we argued above, is very sensitive to the small rolling window samples. Overall, we conclude there is little evidence of rejection of the null of a constant unit root in inflation.

An alternative test is for the hypothesis of parameter stability. The null hypothesis is that the parameters of the autoregression are unchanged for the entire sample period: $\theta_t = \theta$ for all t. Again this is stricter than our null hypothesis since it is conceivable that the scalar

[19] The results are qualitatively similar for the LAR and the half-life, so they are not reported for brevity.

[20] A small issue with these critical values is that they apply to specific lenghts of the rolling and recursive windows. The cutoff for the rolling window is 15 years and a quarter, and 11 years and 1 quarter for the recursive window. We re-estimated the regressions using these window lengths to calculate the test statistics.

functions of persistence remain unchanged despite changes in the parameter vector. Still, parameter stability implies unchanged persistence (even if the converse need not hold) so an acceptance of the null hypothesis implies unchanged persistence. We calculate the sup-Wald or Quandt Likelihood Ratio test, first proposed by Quandt (1960) and studied by Banerjee et al. (1992) when the regressors are possibly non-stationary. The testing procedure consists of computing a sequence of Chow (1960) statistics for parameter stability at each of the different points in time in a middle subset of the sample, and taking the maximum. For our application, the maximum is attained at 1974:3 at a value of 28.69, should be compared with a critical value at 10% significance level of 31.78. We do not reject the null hypothesis of parameter stability and again infer that persistence is best described as unchanged.

6 Comparison with other papers in the literature

Throughout the paper, we have compared our approach with that in Cogley and Sargent (2002). Stock (2002) and Taylor (2000) also preceded us in examining inflation persistence in the United States. Stock (2002) applied a subset of the classical methods in this paper and found no evidence of a change in persistence. Taylor (2000) found that the median-unbiased estimate for the LAR is much lower for the 1982.1-1999.4 sample than the estimate in 1960.2-1979.4. This result is also present in our application—the latter sample yields a LAR estimate of 1.01 and the former 0.64. Yet our 14-year window estimate at 1999.4 is much higher at 1.02, consonant with the earlier sample period. Taylor's results seem therefore to be driven by the anomaly of having the 1982-1983 period at the beginning of the estimation sample.

After the first draft of this paper was written and circulated (in May 2001), a small literature sprung devoted to measuring the persistence of inflation in the United States. Cogley and Sargent (2005) addressed one of our criticisms. They estimated a model in which the variance of innovations can vary over time. However, they did not address the other issues raised in this paper.

Benati (2002) used a classical autoregressive model allowing for time-varying volatility. He estimated it using U.S. data since 1793 and confirmed the Barsky (1987) result that inflation has only been persistent since World War I. Kim et al. (2004) found evidence for a structural break in inflation in late 1979, resulting in lower persistence. However, their application allows for only one break in the sample period, so if there was a very short-lived change in inflation persistence around this date, their test would be unable to detect a return to high levels right after. Our conclusion that persistence is unchanged since 1970 with a possible very short-lived exception in the early 1980s is not inconsistent with their results.

Stock and Watson (2003) agreed with our conclusion. They also found no evidence of a change in persistence in the United States. They further found strong evidence of a fall in volatility. O'Reilly and Whelan (2004) applied some of our methods to the Euro area data since 1970 and also found that persistence was high and unchanged.

An alternative hypothesis has been put forward by Kozicki and Tinsley (2002) and Levin and Piger (2002). They argued that there has been a change in the intercept of the inflation equation (θ_0). Kozicki and Tinsley (2002) interpreted this shift as a change in the long-run inflation target of the Federal Reserve and found some evidence that once one controls for it, inflation persistence was somewhat lower during the 1990s. Levin and Piger (2002) found evidence for a fall in the intercept in 1991, and argue that our 14-year rolling windows might just have missed this shift. After controlling for the shift, they found no evidence that persistence has changed, but their estimates are significantly lower than ours. According to them, persistence is low and unchanged, rather than high and unchanged. The models in our paper allowed for changes in the intercept. The Bayesian model in section 4 includes a time-varying intercept, and the rolling windows in the classical methods in section 5 should at least partially address this concern. However, and in spite of allowing for these shifts, we found that persistence was high and unchanged. One possible explanation for the different results may be that the changes in the intercept occurred suddenly and only a few times and our methods may have little power to detect them.[21] Another explanation, suggested by Levin and Piger (2002), is that in a few years, the extra data points will show a break in 1991. We cannot rule out these possibilities, although we are skeptical.

Our reading of these papers, and acknowledging that others might see things differently, is that the weight of the evidence confirms the main finding in this paper. There is no evidence of a change in persistence in the United States since 1965, with the minor exception of a possible short-lived change during the 1982-83 period. One further robust finding seems to be that there has been a fall in volatility. Finally, there is some evidence suggesting a change in the intercept of the inflation process in 1991, but it still too early to be sure.

7 Conclusion

We have estimated a quarterly time-series for inflation persistence in the United States since 1965. We employed different estimation methods and different measures of persistence. All led to the same conclusion: persistence of inflation has been high and approximately constant over time. The only possible change in persistence refers to a very short period in time (1981-83), an historically exceptional period in monetary policy which should not be seen as a structural change in the economy. Another conclusion is that confidence intervals and credible sets are wide enough that there remains great uncertainty on the exact value of inflation persistence at any given point in time.

Our results provide some answers to the questions raised in the literature described in the introduction. Concerning the dangers of recidivism by monetary policymakers, our results do not exclude this possibility, but its occurrence depends very much on the model used to test the natural rate hypothesis, i.e., the hypothesis that the sum on the coefficients in inflation (or the LAR) is one. Within the Bayesian framework, if the model as exactly implemented by Cogley and Sargent (2002) is used, then the hypothesis is surely rejected - the model

[21] One must be careful not to fall into a tautology here. A stationary model with low persistence but frequent mean shifts is close to undistinguishable from a persistent process or even a unit root. The issue is one of frequency of the changes in long-run expected inflation in response to shocks.

restricts the parameter space and imposes that inflation is always stationary. Without this parameter restriction though, there is no evidence pointing to a rejection of a unit root in inflation. Within the classical framework, if the researcher is careful to compute corrected median unbiased estimates, she will not reject the null. However, if the policymaker applies tests of the natural rate hypothesis exactly as in Solow and Tobin, by using the simple least squares estimates of equations (1) or (8) and testing the null of $\gamma = \rho = 1$, it is possible that she erroneously rejects long-run neutrality.

A second implication of our results is that theories should predict very persistent inflation rates. In regards to the changes in the way monetary policy is conducted, the results show that, even though these likely had effects on the mean or variance of inflation, they seem to have left inflation persistence unchanged.

Finally, the wide confidence intervals and credible sets that we find point to the need to model monetary policy under uncertainty. For most of our estimates, we could not distinguish in the data between LARs of 0.9, 0.95 and 1. Taking the case of an AR(1) for illustration, the first implies a half-life response to shocks of about 1.5 years, the second already 4.25 years, and in the third the shocks persist forever. All are substantively different and likely to lead to very different optimal monetary policies.

Appendix

This appendix describes the details in the implementation of the Bayesian time-varying model. The reader is referred to Cogley and Sargent (2002) for an alternative exposition. The aim is to obtain the posterior:

$$p(\Pi^{T+1,T+H},\theta^{T+1.T+H},\theta^{T},V \quad | \quad \Pi^{T}) =$$
$$p\left(\theta^{T},V \mid \Pi^{T}\right).p(\Pi^{T+1,T+H},\theta^{T+1.T+H} \quad | \quad \Pi^{T},\theta^{T},V).$$

Since analytical expression of the two densities is unavailable we use Monte Carlo methods to simulate them. The algorithm is split in two parts: first, $p\left(\theta^{T},V \mid \Pi^{T}\right)$ is estimated using the Gibbs sampler; and then $p\left(\Pi^{T+1,T+H},\theta^{T+1,T+H} \mid \Pi^{T},\theta^{T},V\right)$ is estimated using the draws $\left(\theta^{T},V\right)$ of step 1 as starting points for future trajectories $\left(\Pi^{T+1,T+H},\theta^{T+1,T+H}\right)$.

Step 1: $\mathbf{p}\left(\theta^{T},V \mid \Pi^{T}\right)$

The following three results are useful::

- $p\left(\theta^{T},V \mid \Pi^{T}\right) \propto p\left(\Pi^{T} \mid \theta^{T},V\right)p\left(\theta^{T},V\right)$ from Bayes rule.

- $p\left(\Pi^{T} \mid \theta^{T},V\right)p\left(\theta^{T},V\right) \propto f\left(\Pi^{T} \mid \theta^{T},V\right)p\left(\theta^{T} \mid V\right)p\left(V\right)$ since the measurement equation is linear and has normal innovations.

- $f\left(\Pi^{T} \mid \theta^{T},V\right)p\left(\theta^{T} \mid V\right)p\left(V\right) \propto I(\theta^{T})\left[f\left(\Pi^{T} \mid \theta^{T},V\right)f\left(\theta^{T} \mid V\right)p\left(V\right)\right]$ using eq. (2). If we did not exclude unstable roots ($I(.)=1$ always), the model would have a linear Gaussian state-space representation with transition equation $f\left(\theta^{T} \mid V\right)$. The posterior kernel would then be given by $p_{L}\left(\theta^{T},V \mid \Pi^{T}\right) \propto f\left(\Pi^{T} \mid \theta^{T},V\right)f\left(\theta^{T} \mid V\right)p\left(V\right)$. We may therefore write $p\left(\theta^{T},V \mid \Pi^{T}\right) \propto I(\theta^{T})p_{L}\left(\theta^{T},V \mid \Pi^{T}\right)$.

The Monte Carlo simulation of $p\left(\theta^{T},V \mid \Pi^{T}\right)$ proceeds in two steps. First, we estimate the unrestricted posterior $p_{L}\left(\theta^{T},V \mid \Pi^{T}\right)$ by simulations and second, we reject the draws that violate the stability condition.

We have also two steps in the estimation of $p_{L}\left(\theta^{T},V \mid \Pi^{T}\right)$. Under some regularity conditions (Roberts and Smith, 1994), the sequence of draws from $p_L\left(\theta^T \mid V, \Pi^T\right)$ and $p_L\left(V \mid \theta^T, \Pi^T\right)$ converge to a draw from the joint distribution $p_L\left(\theta^T, V \mid \Pi^T\right)$. This is the Gibbs sampler and its two steps are the following:

Gibbs Step 1: $p_L\left(\theta^T \mid V, \Pi^T\right)$

Since the unrestricted transition law is linear and normal, the states are also normal: $p_L\left(\theta^T \mid V, \Pi^T\right) = f\left(\theta^T \mid V, \Pi^T\right)$. Moreover, successive applications of Bayes law imply: $f\left(\theta^T \mid V, \Pi^T\right) = f\left(\theta_T \mid V, \Pi^T\right) \prod_{t=1}^{T-1} f\left(\theta_t \mid \theta_{t+1}, V, \Pi^T\right)$.

First, we estimate $f\left(\theta_T \mid V, \Pi^T\right) = N\left(\theta_{T|T}, P_{T|T}\right)$, where $\theta_{T|T}$ and $P_{T|T}$ are given by iterating until T on the Kalman updating equations starting from $\bar{\theta}$ and $\bar{P}$:

$$K_t = \left(P_{t|t-1}X_t + C\right)\left(X_t' P_{t|t-1} X_t + \sigma_\varepsilon^2 + X_t' C + C' X_t\right)^{-1},$$

$$\theta_{t|t} = \theta_{t-1|t-1} + K_t\left(\pi_t - X_t' \theta_{t-1|t-1}\right),$$

$$P_{t|t-1} = P_{t-1|t-1} + Q,$$

$$P_{t|t} = P_{t|t-1} - K_t\left(X_t' P_{t|t-1} + C'\right).$$

Second, we estimate $f\left(\theta_t \mid \theta_{t+1}, V, \Pi^T\right) = N\left(\theta_{t|t+1}, P_{t|t+1}\right)$ by going backwards in time through the sample to obtain $\theta_{t|t+1} = \theta_{t|t} + P_{t|t} P_{t+1|t}^{-1}\left(\theta_{t+1} - \theta_{t|t}\right)$ and $P_{t|t+1} = P_{t|t} - P_{t|t} P_{t+1|t}^{-1} P_{t|t}$. Notice that those are not the equations of the Kalman smoother since the Kalman smoother does not condition on θ_{t+1} to estimate the mean and variance of $f\left(\theta_t \mid \theta_{t+1}, V, \Pi^T\right)$.

Gibbs Step 2: $p_L\left(V \mid \theta^T, \Pi^T\right)$

The inverse Wishart prior on the hyperparameters combined with a Gaussian likelihood yields an inverse Wishart posterior given by: $p_L\left(V \mid \theta^T, \Pi^T\right) = IW(V_1^{-1}, T_1)$, where $T_1 = T_0 + T$, $V_1 = \bar{V} + \bar{V}_T$ and $\frac{1}{T}\bar{V}_T = \frac{1}{T}\sum_{t=1}^{T}\left(\varepsilon_t v_t\right)\left(\varepsilon_t v_t\right)'$.

To sample from an inverse Wishart, we use two facts: First, that $p_L\left(V^{-1} \mid \theta^T, \Pi^T\right) = W(V_1^{-1}, T_1)$. Second, that sampling from $W(V_1^{-1}, T_1)$ is equivalent to taking T_1 independent draws η_i from a normal density $N\left(0, V_1^{-1}\right)$, computing $V^{-1} = \sum_{i=1}^{T_1} \eta_i \eta_i'$ and inverting the result.

Rejection sampling:

To impose the $I(.)$ prior restrictions, we check the autoregressive roots of each simulation at each date and reject non-stationary representations.

Step 2: $p\left(\Pi^{T+1,T+H},\theta^{T+1,T+H} \mid \Pi^{T},\theta^{T},V\right)$

For every draw of the past (θ^{T},V), we compute the future trajectories. To do so, we factorize:

$$p(\Pi^{T+1,T+H},\theta^{T+1.T+H},\theta^{T},V \quad | \quad \Pi^{T},\theta^{T},V) =$$
$$p\left(\theta^{T+1,T+H} \mid \Pi^{T},\theta_{T},V\right) p(\Pi^{T+1,T+H} \quad | \quad \theta^{T+1,T+H},\Pi^{T},\theta_{T},V),$$

and obtain draws from each of the two components in turn.

Future Step 1: $p\left(\theta^{T+1,T+H} \mid \Pi^{T},\theta_{T},V\right)$

Since the states are Markov processes, we have that: $p\left(\theta^{T+1,T+H} \mid \Pi^{T},\theta_{T},V\right) = \prod_{i=1}^{H} p\left(\theta_{T+i} \mid \Pi^{T},\theta_{T+i-1},V\right)$. We may then use the fact that $\theta_{t+i} = \theta_{t+i-1} + v_{t+i}$, where $v_{t+i} \overset{iid}{\sim} N(0,Q)$. We take H draws from $N(0,Q)$ and iterate on the state equation to obtain a future trajectory. As before, when applicable we also reject the explosive roots associated with each draw.

Future Step 2: $p\left(\Pi^{T+1,T+H} \mid \theta^{T+1,T+H},\Pi^{T},\theta_{T},V\right)$

This joint distribution may be factored into: $p\left(\Pi^{T+1,T+H} \mid \theta^{T+1,T+H},\Pi^{T},\theta_{T},V\right) = \prod_{i=1}^{H} p\left(\pi_{T+i} \mid \Pi^{T+1,i-1},\theta^{T+1,T+H},\theta_{T},V,\Pi^{T}\right)$. But then, using the fact that $\varepsilon_{t+i} \sim N\left(C'Q^{-1}v_{t+i}\ ,\ R-C'Q^{-1}C\right)$, it follows that the distribution of inflation is $\pi_{T+i} \sim N\left(X'_{T+1}\theta_{T+1} + C'Q^{-1}v_{t+i}\ ,\ R-C'Q^{-1}C\right)$. We take draws for ε_{t+i} and the lags of π_{T+i} in the data to iterate the measurement equation $\pi_{t+i} = X'_{t+i}\theta_{t+i} + \varepsilon_{t+i}$.

This gives the future draws for inflation that we use to obtain our measures of persistence.

References

Andrews, D., 1993. Exactly median-unbiased estimation of first order autoregressive / unit root models. Econometrica 61 (1), 139-165.

Andrews, D., Chen, H.-Y., 1994. Approximately median-unbiased estimation of autoregressive models. Journal of Business and Economic Statistics 12 (2), 187-204.

Banerjee, A., Lumsdaine, R., Stock J., 1992. Recursive and sequential tests of the unit-root and trend-break hypotheses: theory and international evidence. Journal of Business and Economic Statistics 10 (3), 271-288.

Barsky, R., 1987. The Fisher hypothesis and the forecastibility and persistence of inflation. Journal of Monetary Economics 19, 3-24.

Benati, L., 2002. Investigating inflation persistence across monetary regimes. Bank of England working paper.

Boivin, J., Giannoni, M., 2005. Has monetary policy become more effective? Review of Economics and Statistics, forthcoming.

Brainard, W., Perry, G., 2000..Making policy in a changing world, in: Perry, G., Tobin J. (Eds.), Economic Events, Ideas, and Policies: The 1960s and After. Brookings Institution Press, Washington.

Chow, G., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28, 591-605.

Clarida, R., Gali, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. Quarterly Journal of Economics 115 (1), 147-180.

Cogley, T., Sargent, T., 2002. Evolving post-World War II U.S. Inflation Dynamics, in: Gertler M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2001. MIT Press: Cambridge.

Cogley, T., Sargent, T., 2005. Drifts and volatilities: monetary policies and outcomes in the post World War II U.S. Review of Economic Dynamics 8 (2), 262-302.

DeJong, D., Whiteman, C., 1991. Reconsidering 'Trends and random walks in macroeconomic time series'. Journal of Monetary Economics 28, 221-254.

Dickey, D., Fuller, W., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74 (366), 427-431.

Hall, R., 1999. Comment on: Rethinking the role of the NAIRU in monetary policy: implications of model formulation and uncertainty, in: Taylor, J. (Ed.), Monetary Policy Rules. NBER and University of Chicago Press: Chicago, Illinois.

Kendall, M., Stuart, A., 1967. The Advanced Theory of Statistics, vol. 2, 2nd edition. Charles Griffin: London.

Kim, C., Nelson, C., Piger, J., 2004. The less volatile U.S. economy: a Bayesian investigation of timing, breadth, and potential explanations. Journal of Business and Economics Statistics 22 (1), 80-93.

Kozicki, S., Tinsley, P., 2002. Alternative sources of the lag dynamics of inflation. Federal Reserve Bank of Kansas City Working Paper 02-12.

Levin, A., Piger, J., 2002. Is inflation persistence intrinsic in industrial economics? Federal Reserve Bank of Saint Louis Working Paper 2002-023.

McConnell, M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980's? American Economic Review 90 (5), 1464-1476.

O'Reilly, G., Whelan, K., 2004. Has Euro-area inflation persistence changed over time? European Central Bank Working Paper 335.

Phillips, P., 1991. To criticize the critics: an objective Bayesian analysis of stochastic trends. Journal of Applied Econometrics 6 (4), 333-364.

Quandt, R., 1960. Tests of the hypothesis that a linear regression system obeys two separate regimes. Journal of the American Statistical Association 55, 324-330.

Roberts, G., Smith, A. 1994. Some simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. Stochastic Processes and their Applications 49, 207-216.

Rogoff, K., 1996. The purchasing power parity puzzle. Journal of Economic Literature 34, 647-668.

Rossi, B., 2005. Confidence intervals for half-life deviations from purchasing power parity. Journal of Business and Economic Statistics 23 (4), 432-442.

Sargent, T., 1971. A note on the accelerationist controversy. Journal of Money, Credit and Banking 8 (3), 721-725.

Sargent, T., 1999. The Conquest of American Inflation. Princeton University Press: Princeton.

Sims, C., 2002. Comment on: Evolving post World War II U.S. inflation dynamics. In: Gertler, M., Rogoff, K., (Eds.), NBER Macroeconomics Annual 2001. MIT Press: Cambridge.

Solow, R., 1968. Recent controversy on the theory of inflation: an eclectic view. In: Rousseaus, S. (Ed.), Proceedings of a Symposium on Inflation: Its Causes, Consequences and Control. New York University: New York.

Stock, J., 1991. Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series. Journal of Monetary Economics 28 (3), 435-460.

Stock, J., 2002. Comment on: Evolving post-World War II U.S. inflation dynamics. In: Gertler, M., Rogoff, K., (Eds.), NBER Macroeconomics Annual 2001. MIT Press: Cambridge.

Stock, J., Watson, M., 2003. Has the business cycle changed and why? In: Gertler, M., Rogoff, K., (Eds.), NBER Macroeconomics Annual 2002. MIT Press: Cambridge.

Taylor, J., 1998. Monetary policy guidelines for unemployment and inflation stability. In: Friedman, B., Solow, R., (Eds.), Inflation, Unemployment and Monetary Policy, MIT Press: Cambridge.

Taylor, J., 1999. Staggered wage and price setting in macroeconomics. In: Taylor, J., Woodford, M., (Eds.), Handbook of Macroeconomics, vol. 1B. Elsevier North Holland: Amsterdam.

Taylor, J., 2000. Low inflation, pass-through, and the pricing power of firms. European

Economic Review 44 (7), 1389-1408.

Tobin, J., 1968. Discussion. In Rousseaus, S., (Ed.), Procedings of a Symposium on Inflation: Its Causes, Consequences and Control, New York University: New York.

Uhlig, H., 1994. What macroeconomists should know about unit roots: a Bayesian perspective. Econometric Theory 10 (3/4), 645-671.

Table 1.
Summary statistics for inflation over the decades

Decade	*1950s*	*1960s*	*1970s*	*1980s*	*1990s*
Mean	2.47	2.47	6.50	4.47	2.22
Standard Deviation	2.93	1.55	2.10	2.36	0.96
Serial correlation	0.31	0.81	0.68	0.91	0.68

Figure 1. Inflation in the United States.

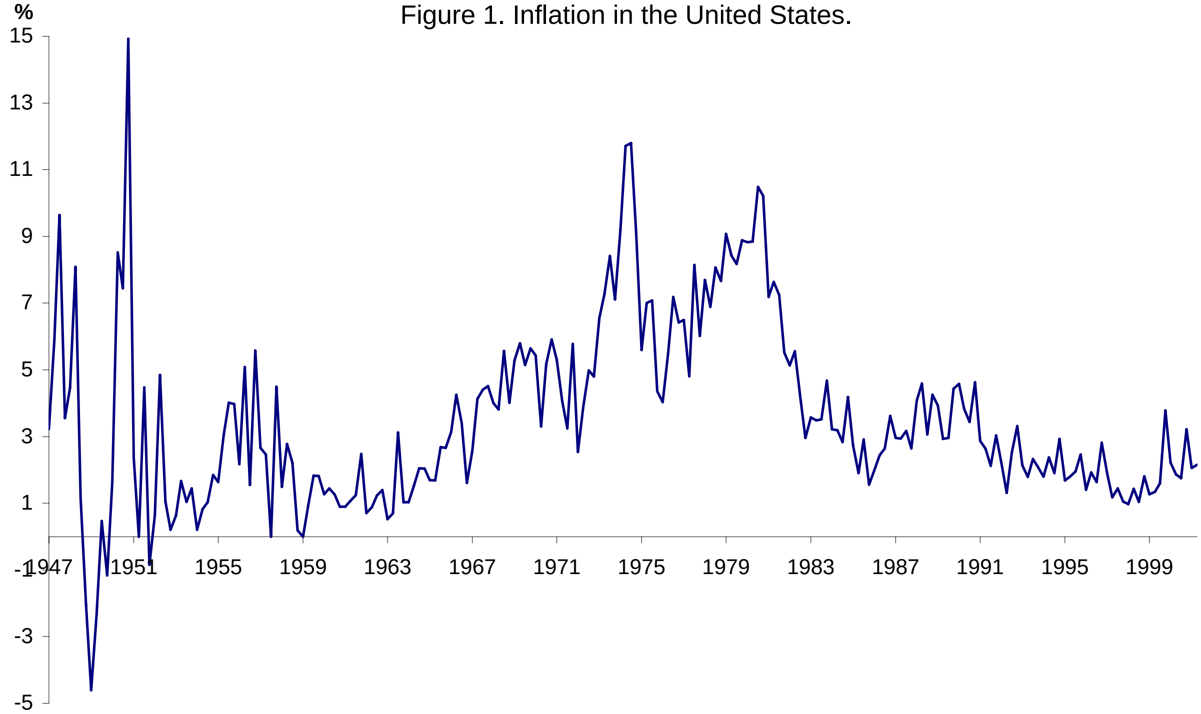


Figure 2a. First-order serial correlation of inflation over 14-year rolling window samples.

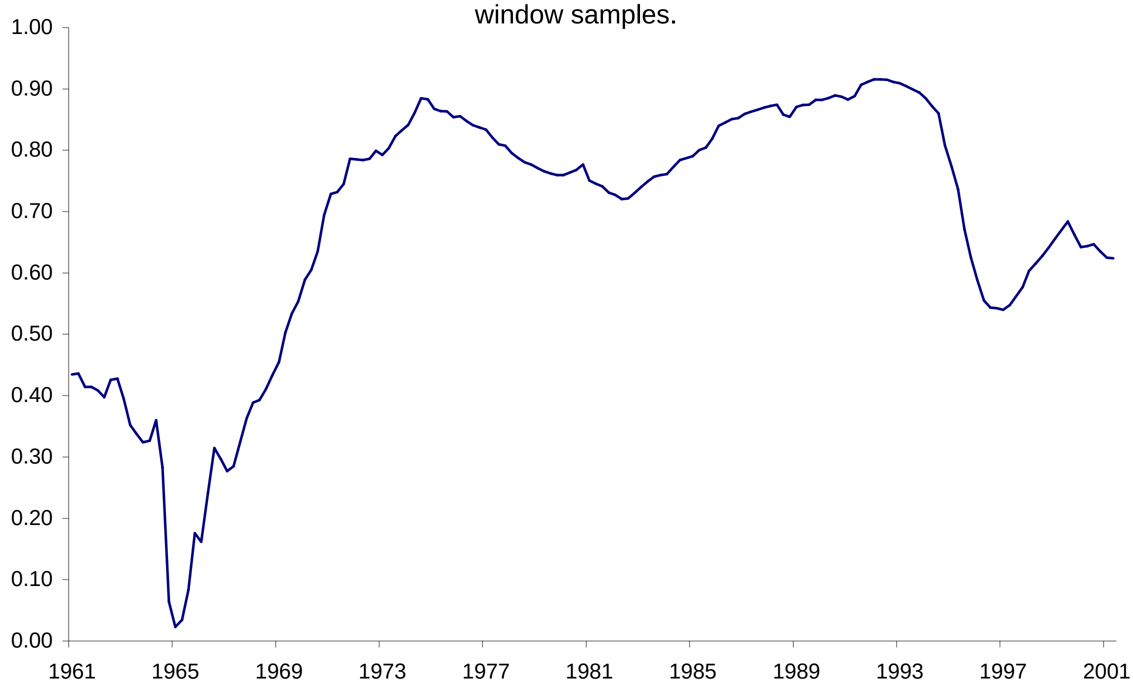
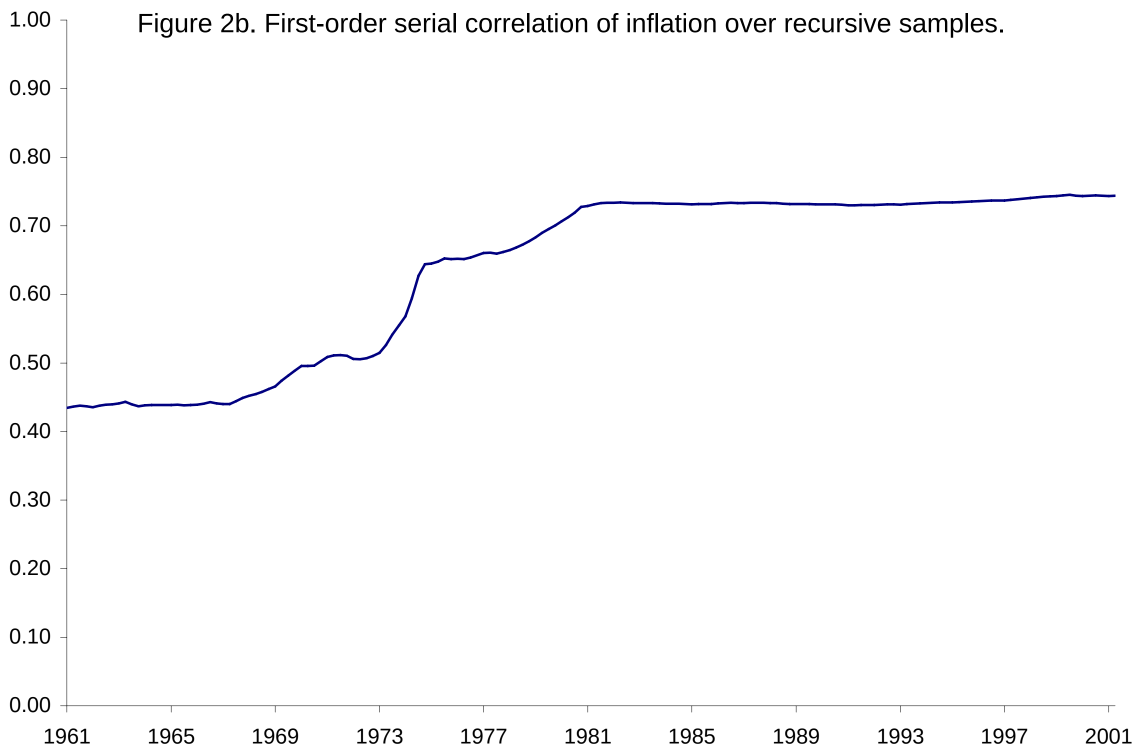


Figure 2b. First-order serial correlation of inflation over recursive samples.



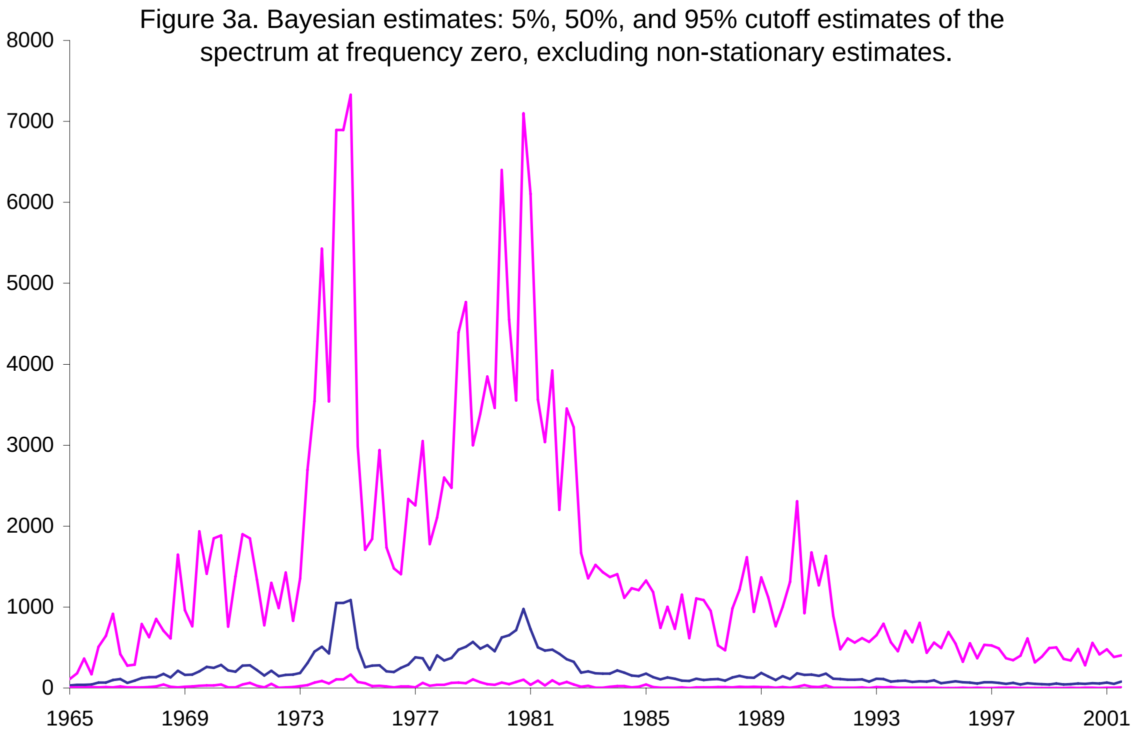


Figure 3a. Bayesian estimates: 5%, 50%, and 95% cutoff estimates of the spectrum at frequency zero, excluding non-stationary estimates.
8000
7000
6000
5000
4000
3000
2000
1000
0
1965
1969
1973
1977
1981
1985
1989
1993
1997
2001

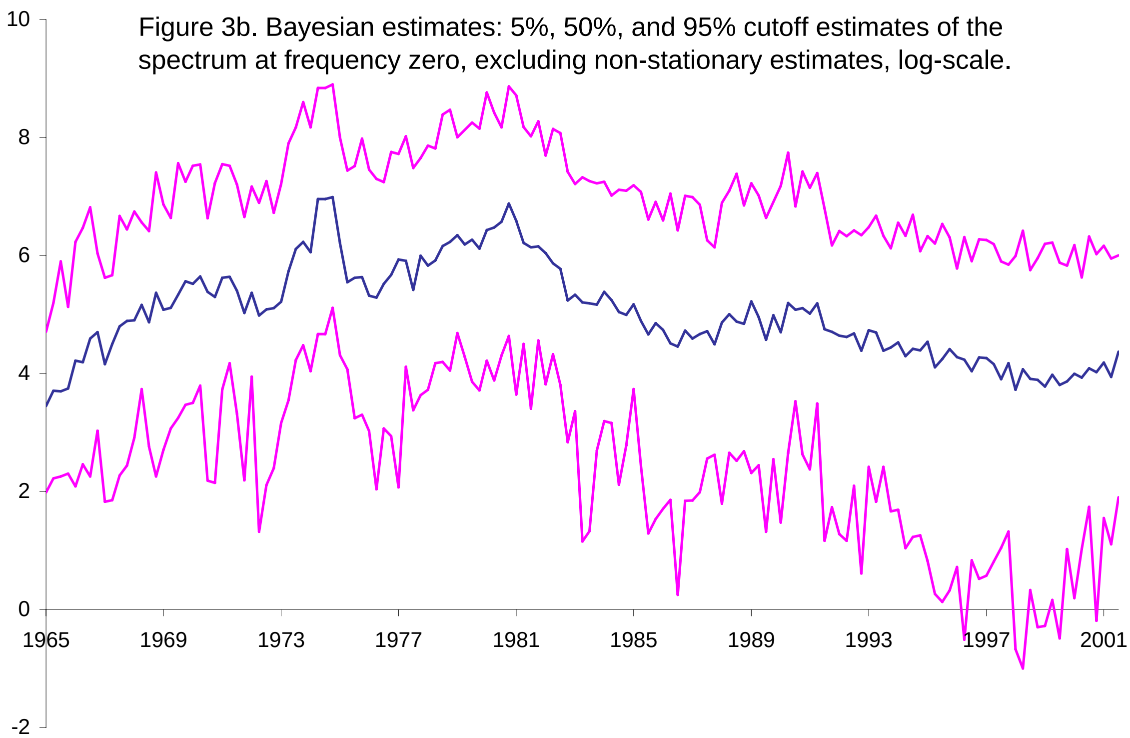


Figure 3b. Bayesian estimates: 5%, 50%, and 95% cutoff estimates of the spectrum at frequency zero, excluding non-stationary estimates, log-scale.
10
8
6
4
2
0
-2
1965
1969
1973
1977
1981
1985
1989
1993
1997
2001

Figure 4a. Bayesian estimates: 5%, 50%, and 95% cutoff estimates of the LAR, excluding non-stationary estimates.

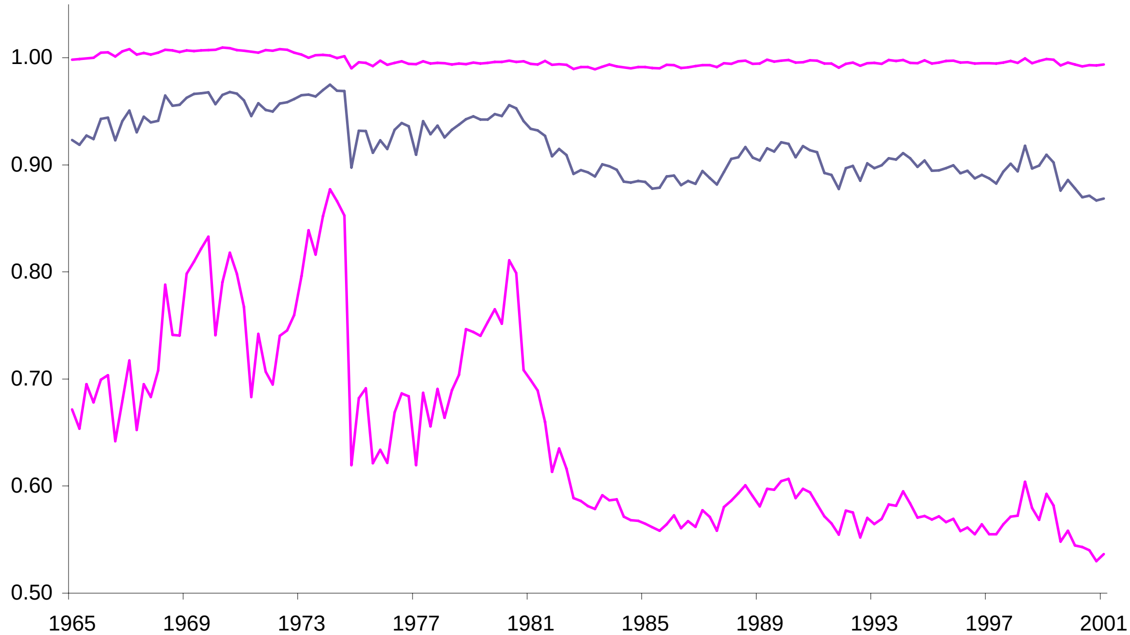


Figure 4b. Bayesian estimates: 5%, 50%, and 95% cutoff estimates of the SUM excluding non-stationary estimates.

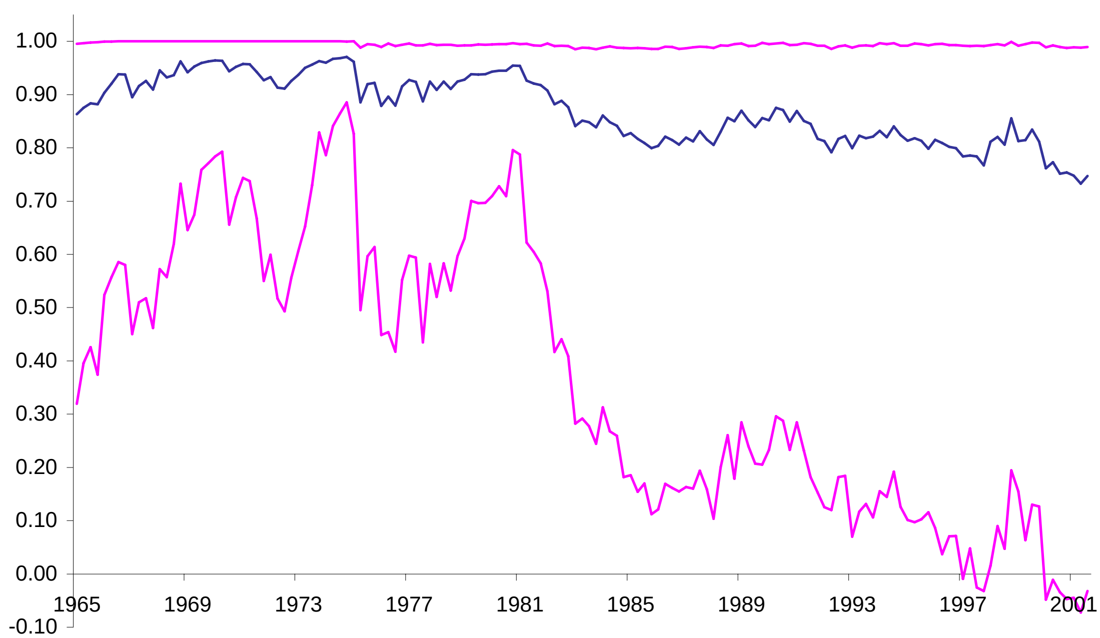


Figure 4c. Bayesian estimates: 5%, 50%, and 95% cutoff estimates of the HL, excluding non-stationary estimates.

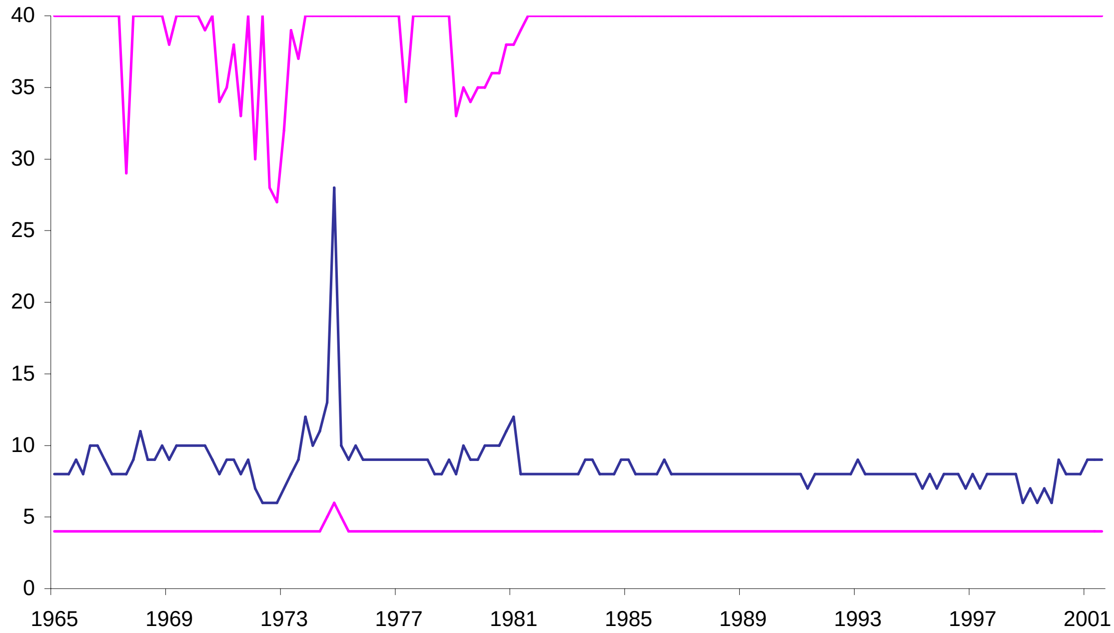


Figure 5a. Bayesian estimates: Mean, 5%, 50%, and 95% cutoff estimates of the LAR.

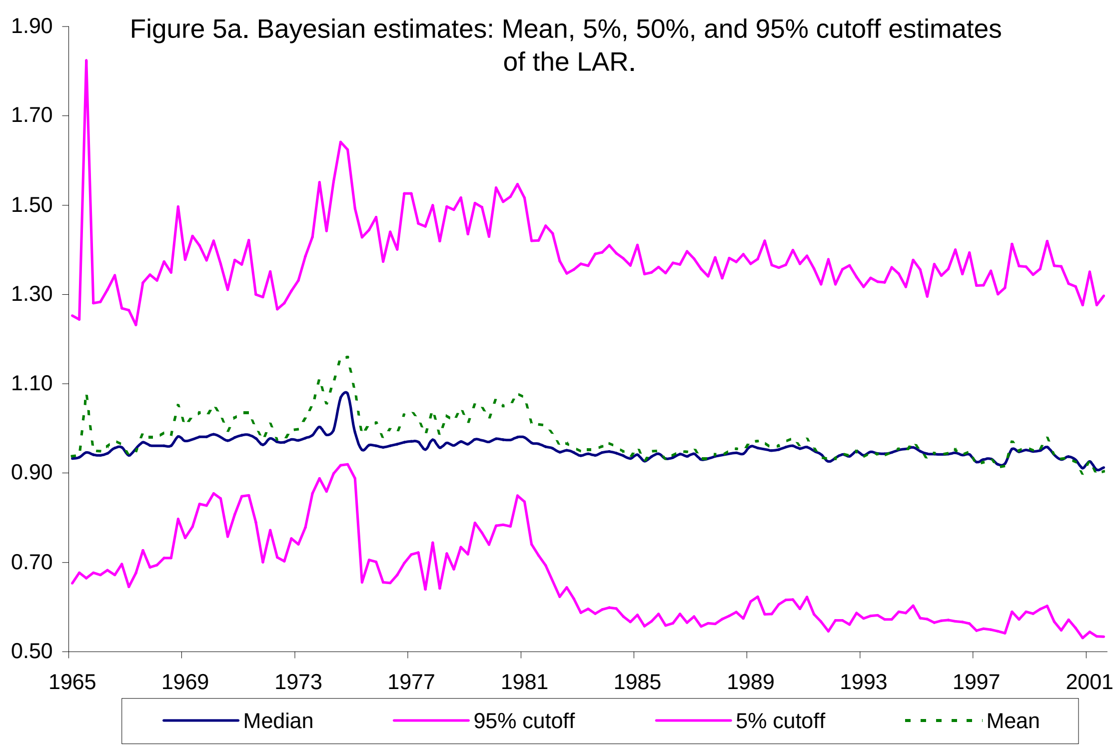


Figure 5b. Bayesian estimates: Mean, 5%, 50%, and 95% cutoff estimates of the SUM.

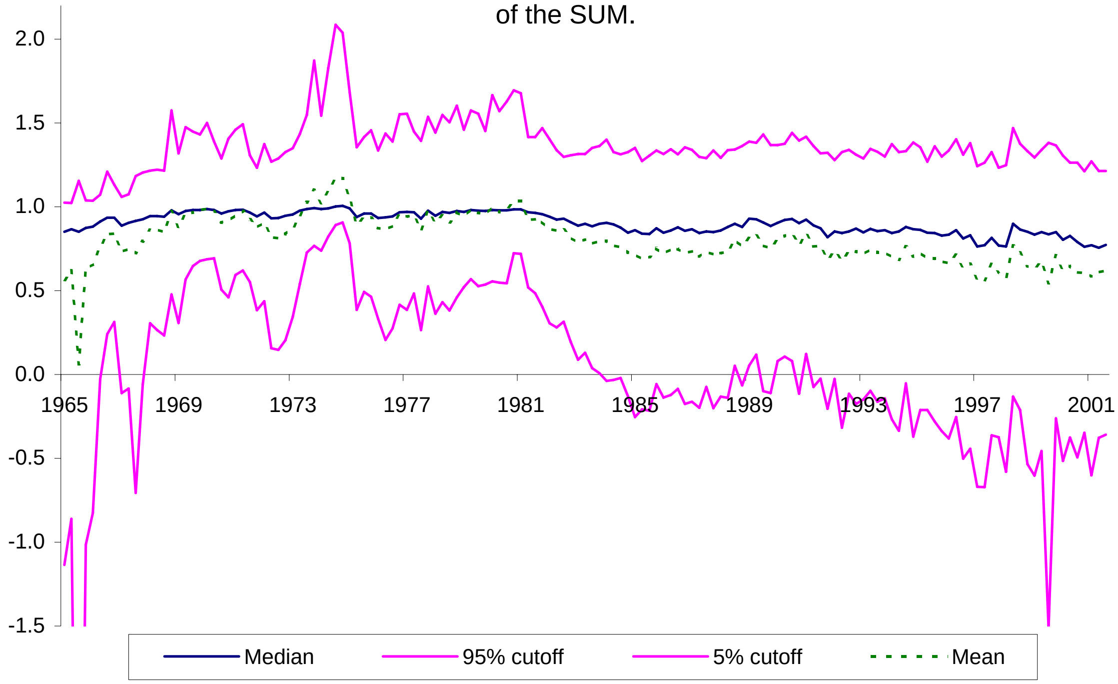


Figure 5c. Bayesian estimates: Mean, 5%, 50%, and 95% cutoff estimates of the HL.

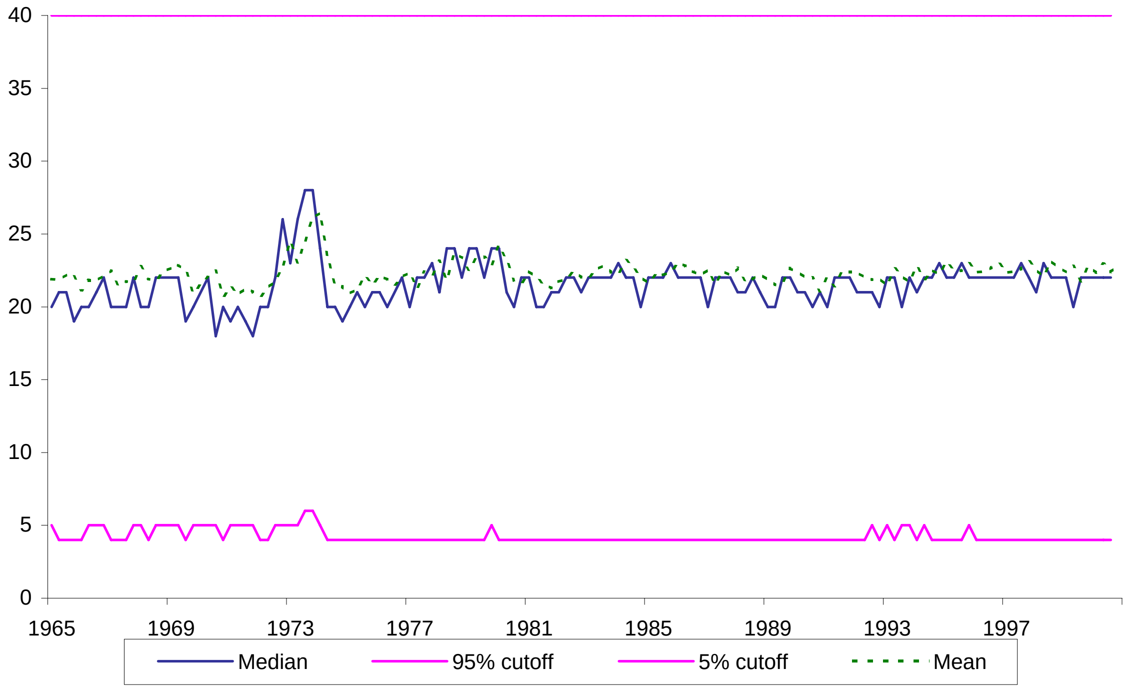


Figure 6a. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the LAR.

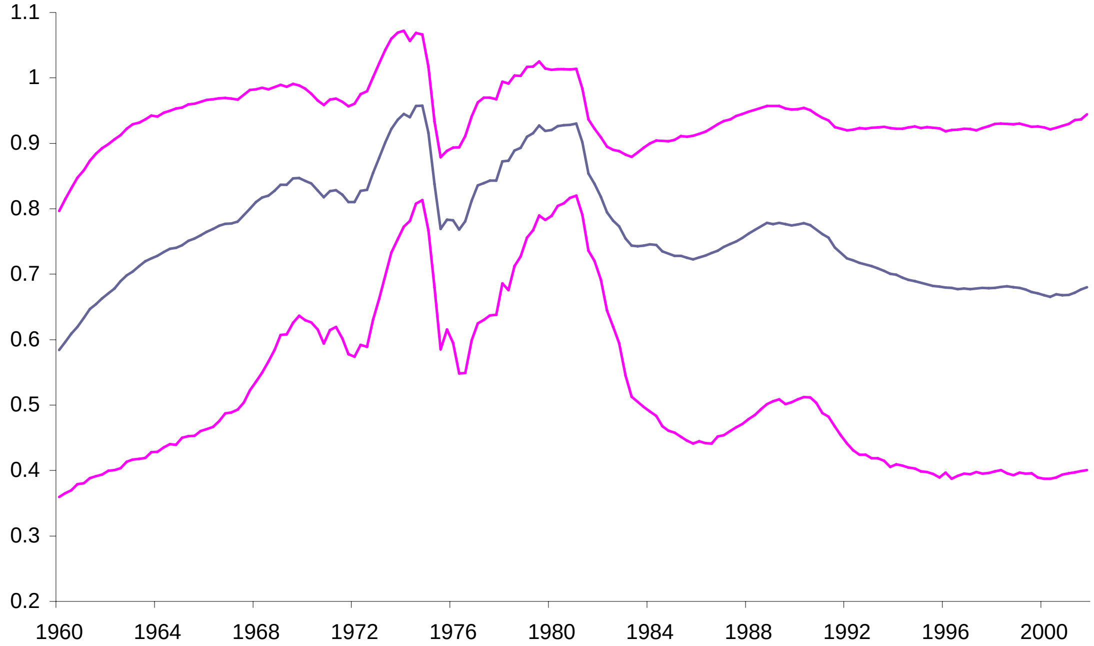


Figure 6b. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the SUM.

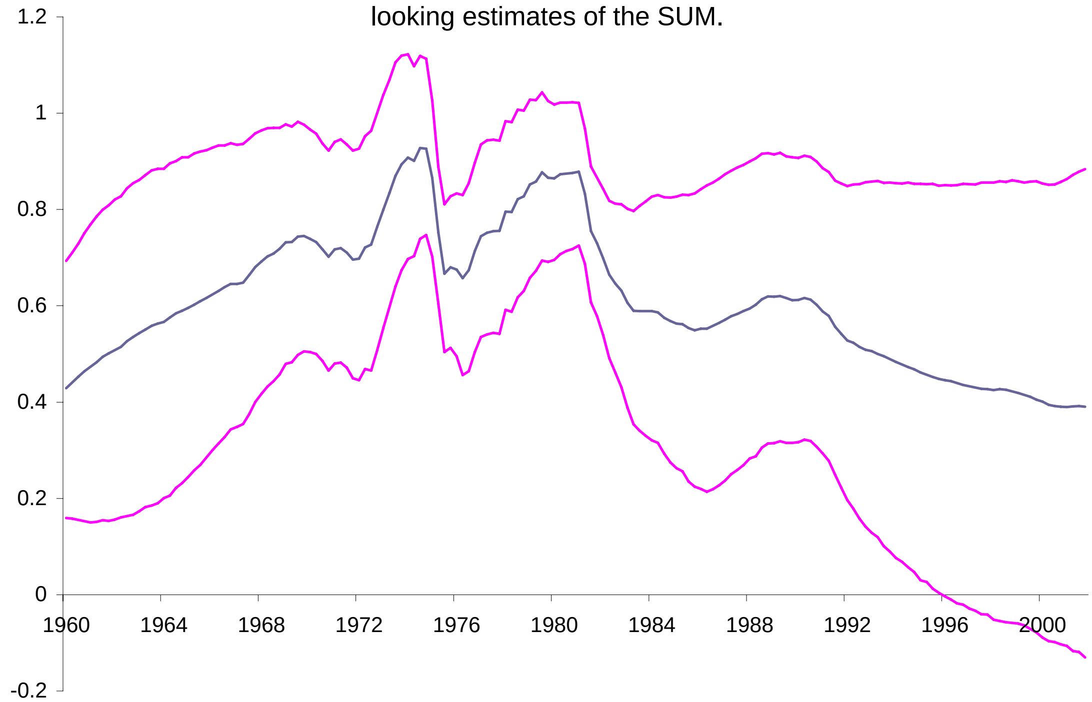


Figure 6c. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the HL.

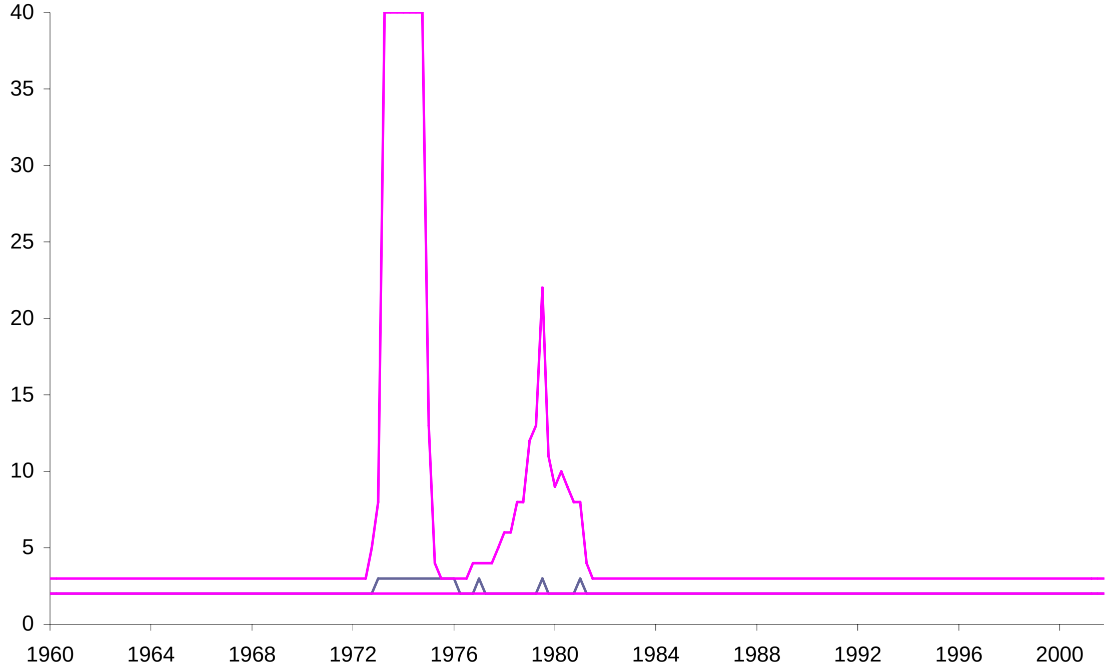


Figure 7a. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the LAR, with alternative priors

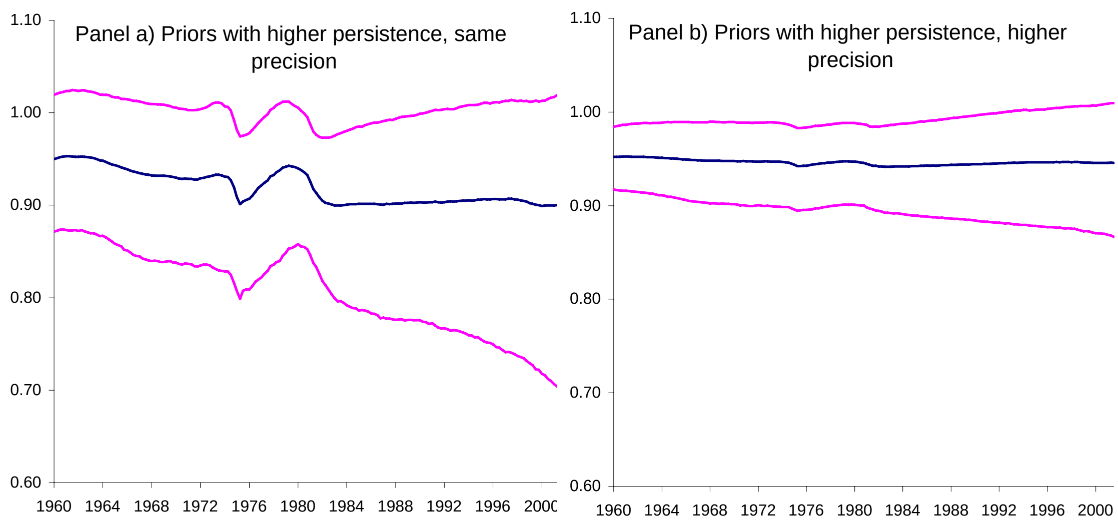


Figure 7b. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the SUM, with alternative priors

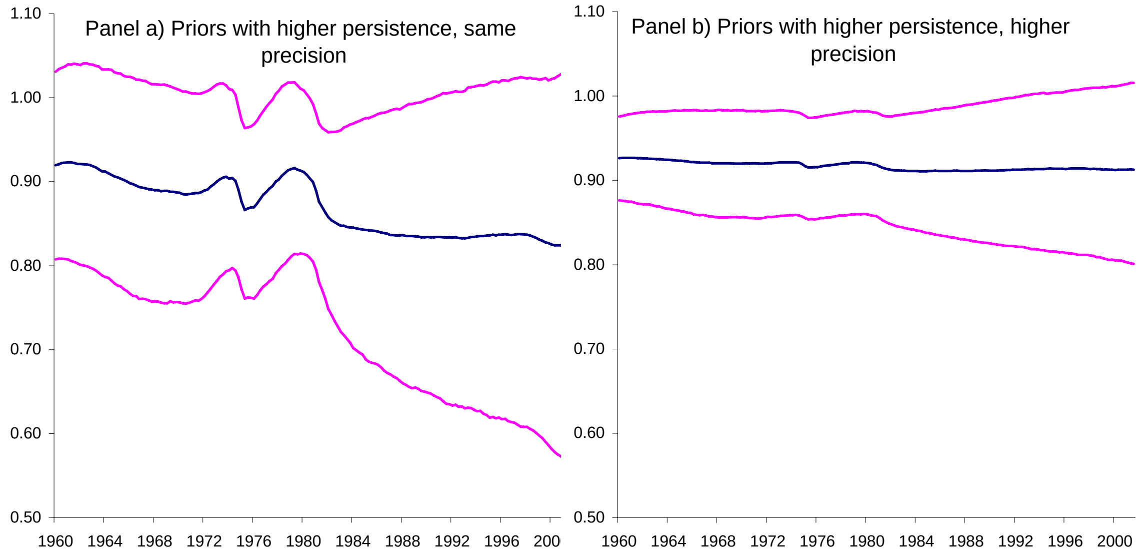


Figure 7c. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the HL, with alternative priors

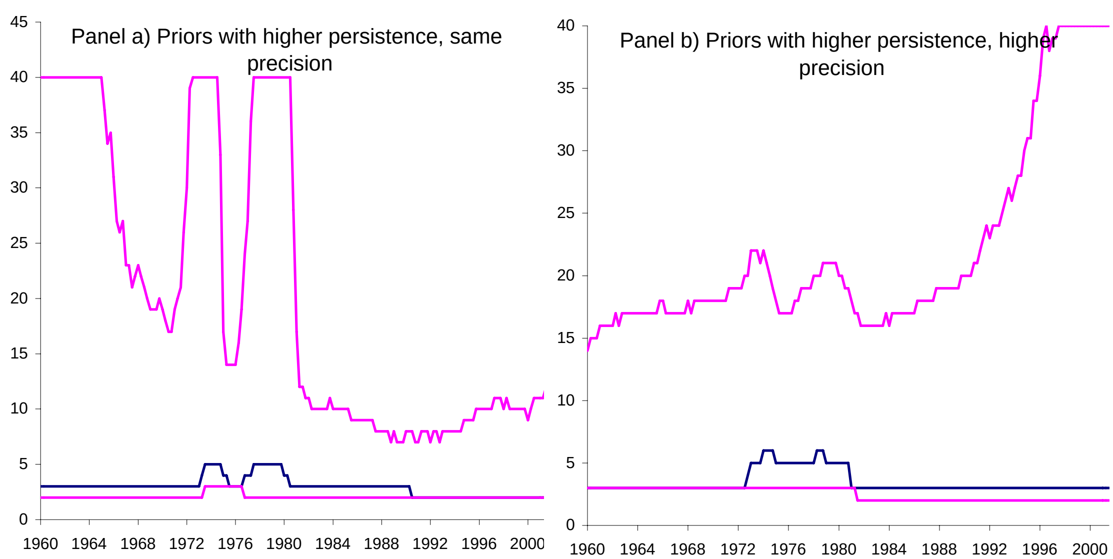


Figure 8. Bayesian estimates: 5%, 50%, and 95% cutoff backward-looking estimates of the LAR, with different number of draws.

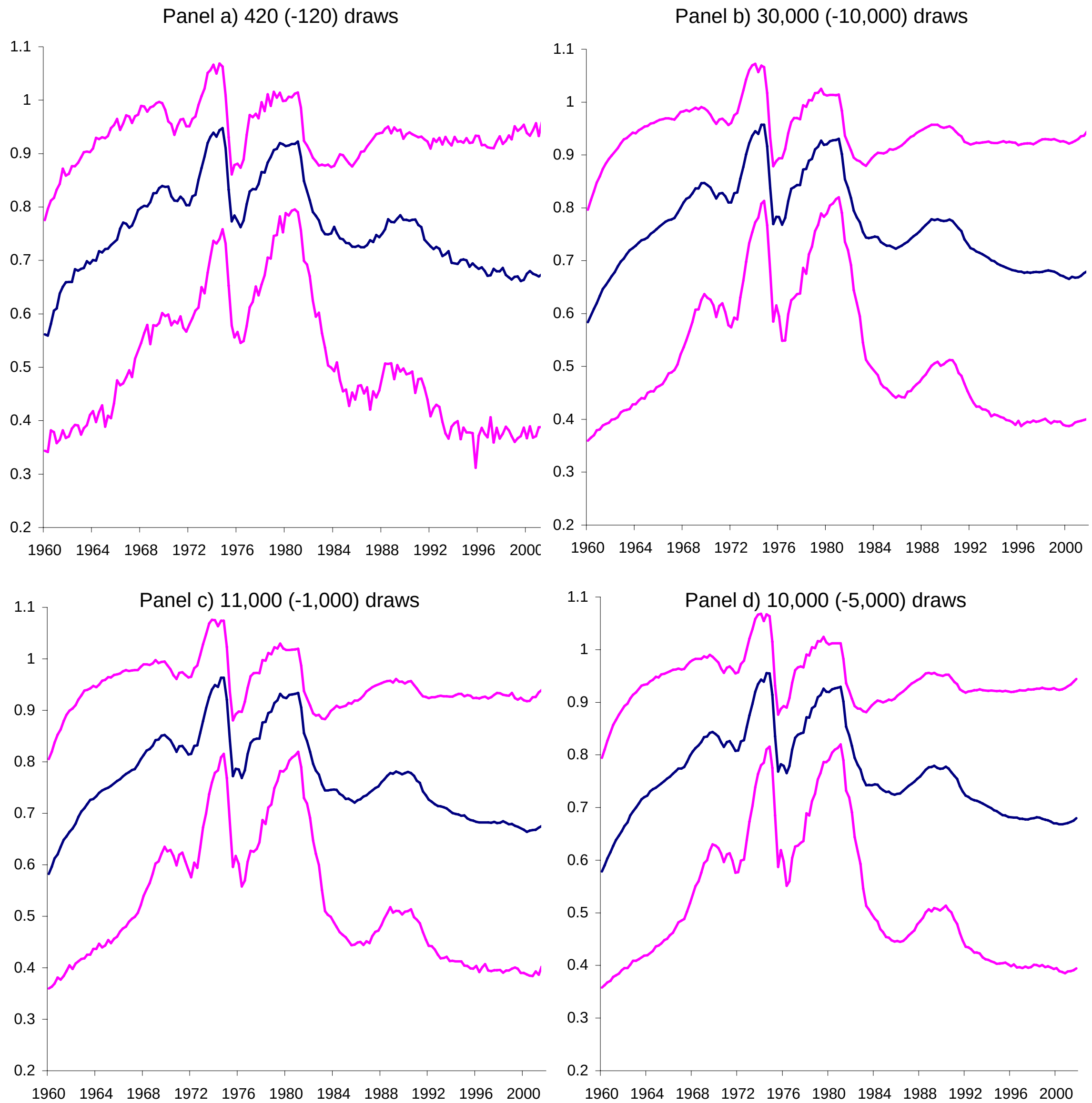


Figure 9a. Classical estimates: median-unbiased and 90% confidence interval estimates of the LAR, over recursive samples.

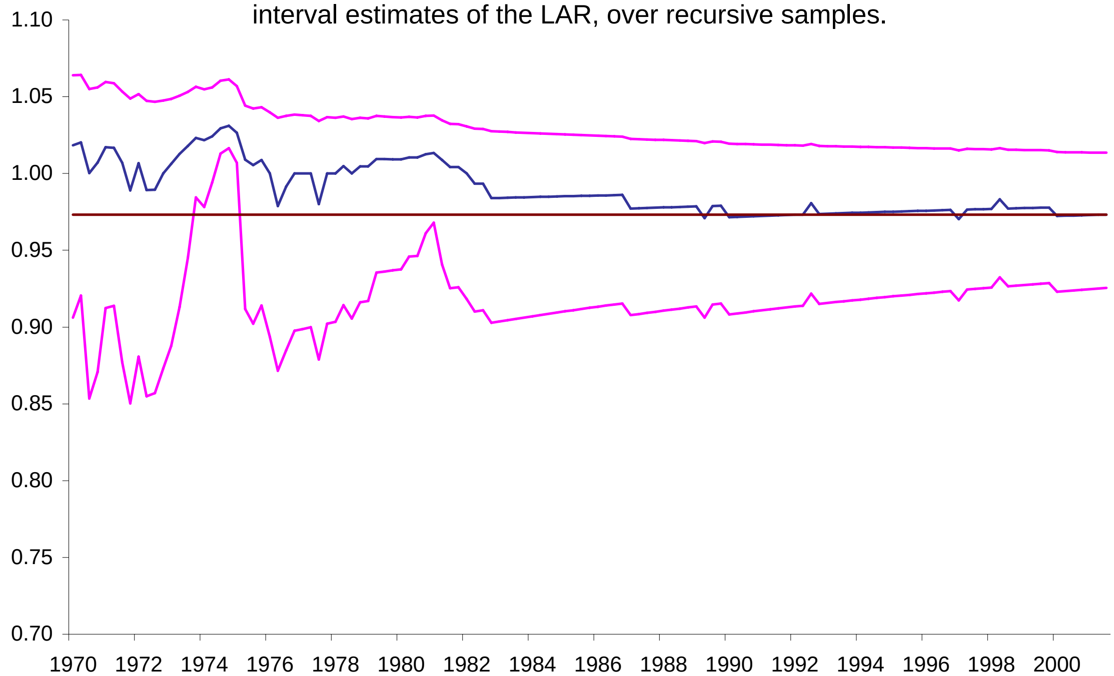


Figure 9b. Classical estimates: median-unbiased and 90% confidence interval estimates of the SUM, over recursive samples.

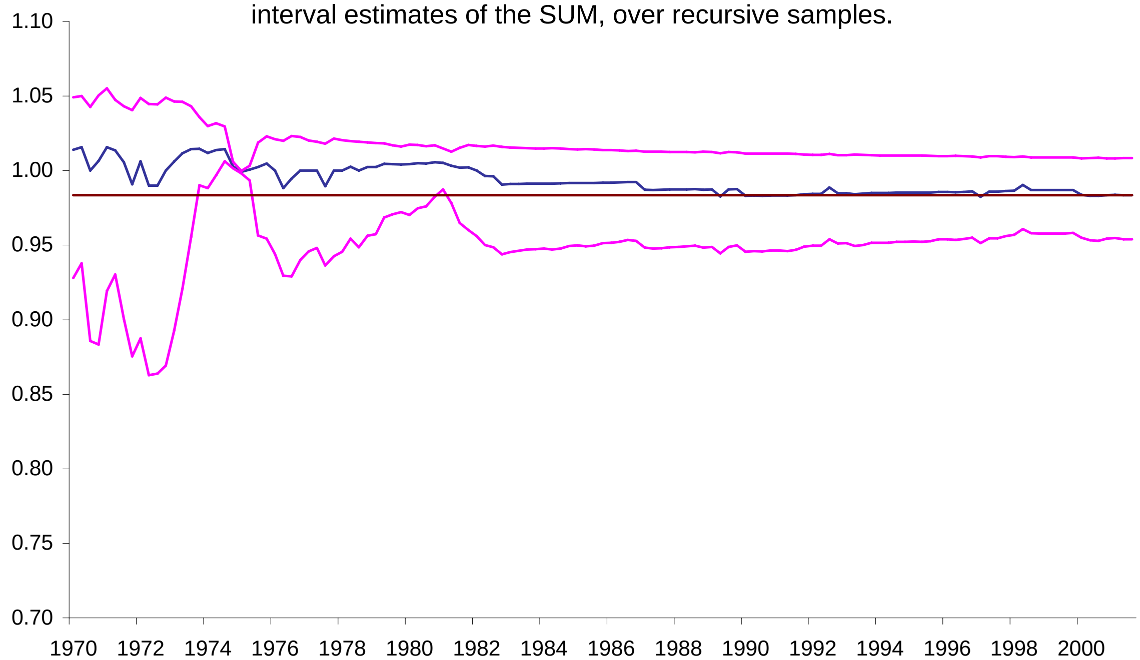


Figure 10a. Classical estimates: median-unbiased and 90% confidence interval estimates of the LAR, over 14-year rolling window samples.

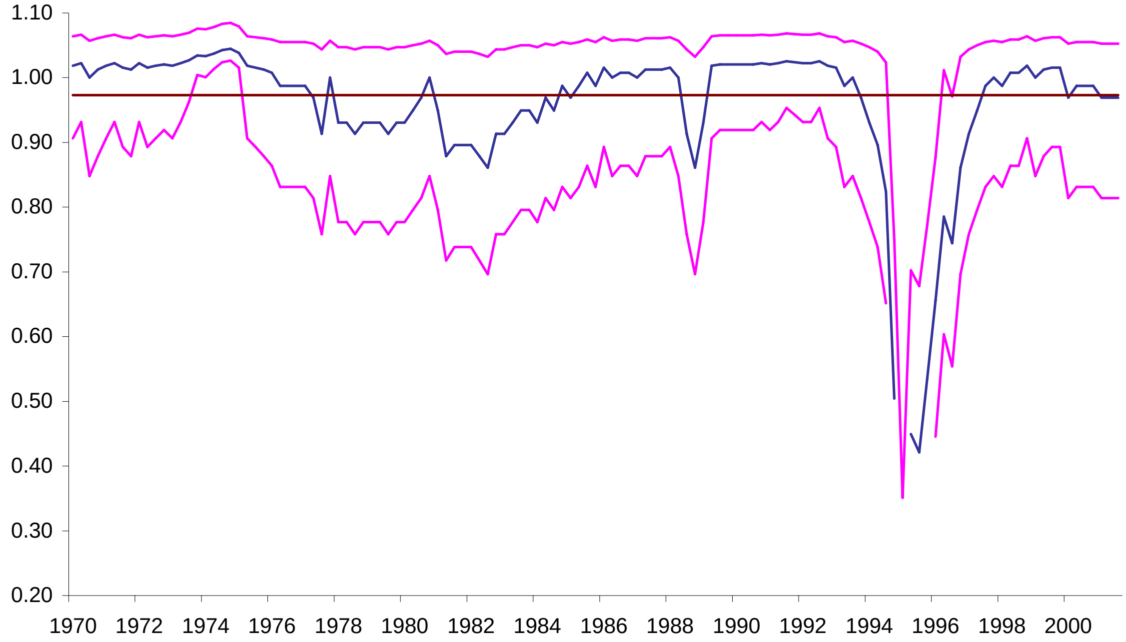


Figure 10a. Classical estimates: median-unbiased and 90% confidence interval estimates of the SUM, over 14-year rolling window samples.

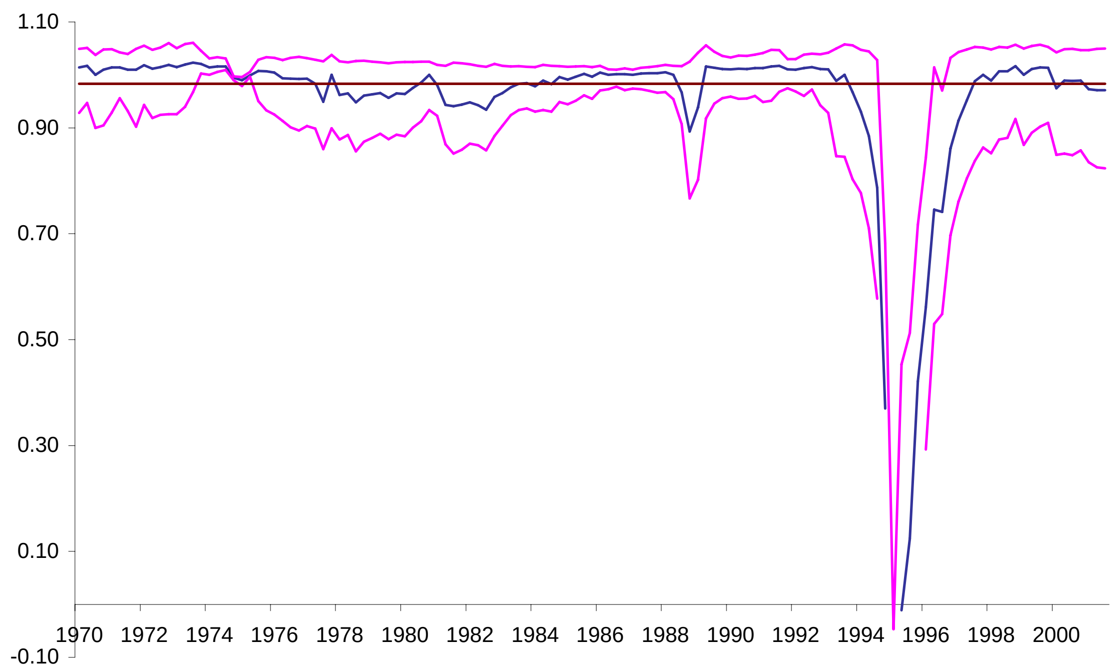


Figure 11a. Classical estimates: approximately median-unbiased and 90% confidence interval estimates of the SUM, over recursive samples.

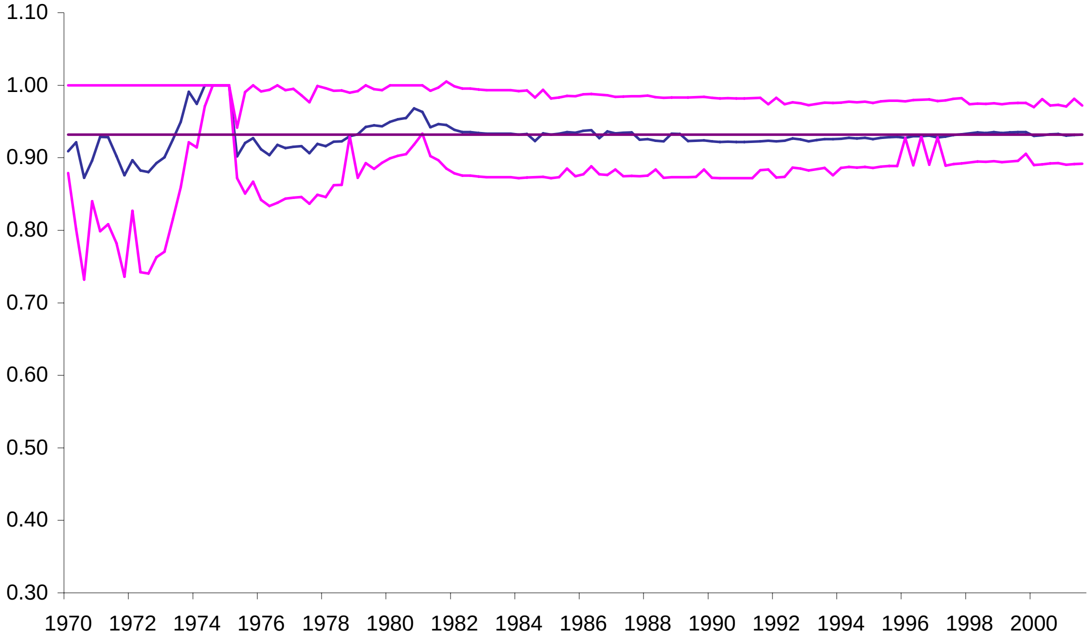


Figure 11b. Classical estimates: approximately median-unbiased and 90% confidence interval estimates of the SUM, over 14-year rolling window samples.