Using mean reversion as a measure of persistence∗

Daniel A. Dias† and Carlos Robalo Marques‡

June 23, 2009

Abstract

This paper suggests a new scalar measure of persistence together with a companion estimator, which has the advantage of not requiring the specification and estimation of a model for the series under investigation. The statistical properties of the companion estimator are established, which allow tests of hypotheses to be performed, under very general conditions. The use of the new measure is illustrated by re-evaluating persistence of inflation for the United States and the Euro Area. The conclusions for the United States do not differ significantly from what has been found in previous empirical studies. However, for the Euro Area we find evidence of a significant break occurring in 2001/2002, such that persistence becomes virtually nil for the period that follows the launch of the euro and the implementation of a common monetary policy by the European Central Bank.

JEL classification codes: E31, C22, E52.

Key Words: Persistence, mean-reversion, non-parametric estimator.

∗We especially thank for useful discussions, Stephen Cecchetti, Matteo Ciccarelli, Jordi Gali, Vitor Gaspar, Andrew Levin, James Stock and participants at the ECB’s Inflation Persistence Network Conference, held in Frankfurt, at the LACEA-LAMES 2007 Conference, held in Bogotá and at the Royal Economic Society Conference 2008, held in Coventry. Helpful suggestions from Patrik Guggenberger, João Santos Silva, José Maria Brandão de Brito and José Ferreira Machado are also acknowledged. The usual disclaimer applies.

†Anderson Graduate School of Management - UCLA and CEMAPRE. Corresponding author. Tel.: +1 310 825 8207. E-mail: daniel.dias.2010@anderson.ucla.edu.

‡Banco de Portugal, Research Department. Tel.: +351 213128330; E-mail: cmrmarques@bportugal.pt.
1 Introduction

Understanding inflation persistence is crucial for central banks because it may have strong implications for the design and implementation of monetary policy. In particular, it may be argued that the appropriate response to shocks hitting the economy depends on the degree to which their effect on inflation is persistent. Furthermore, the horizon at which monetary policy should aim for price stability also depends on the persistence of inflation: with less persistence, inflation can be stabilised in shorter time following a shock, so that the degree of inflation persistence may also be seen as an important factor determining the medium-term orientation of monetary policy. On the other hand, it may also be claimed that persistence of inflation is a major determinant of the economic costs of disinflation\(^1\). No wonder thus, that inflation persistence has been over the last decade an intensely investigated topic in macroeconomics. Issues such as whether inflation is highly persistent or not and its implications for monetary policy strategy, whether it has changed over time or remained constant, whether it is structural or may vary according to the specific monetary policy regime, are examples of relevant questions that have been addressed in the literature.

In a different context some literature has also found important to investigate persistence of other macroeconomic variables, such as aggregate output or the deviations of the economy from purchasing power parity (PPP) conditions.


to name but a few, are examples of contributions to the measurement of persistence of PPP deviations from equilibrium.

In most of the papers that try to evaluate persistence of inflation the "sum of the autoregressive coefficients" emerges as the most popular scalar measure of persistence, while the "half-life" is very popular in the literature that investigates persistence of PPP deviations. Other scalar measures of persistence also used in the literature include, for instance, the "largest autoregressive root" (see Stock, 1991, 2001) and the "spectrum at zero frequency" (see, for instance, Andrews and Chen, 1994). The usefulness of scalar measures of persistence stems from the fact that they are summary measures of the information contained in the impulse response functions of the estimated models. Of course the use of such scalar measures of persistence may be criticised exactly on the grounds that they are not capable of retaining all the potentially relevant features of the underlying impulse response functions. However, by itself, the impulse response function being an infinite-length vector is not very useful as a measure of persistence, especially so if the purpose is to quantify and compare the degree of persistence across different time series. Thus, despite their limitations, the scalar measures of persistence remain a useful way of quantifying persistence of time series data in empirical applications.

All the above cited measures of persistence share the common feature that they are parametric in the sense that are defined and computed by estimating a times series model (usually an autoregressive process) for the data under investigation. This paper contributes for this strand of the literature by suggesting a new measure of persistence, which is broader in scope than the widely used "sum of the autoregressive coefficients" and has the property of being model free, as its use does not require the specification and estimation of a model for the data. This new measure of persistence, denoted in the paper by $\gamma$, relies on the idea that there is a relationship between persistence and mean reversion, and is defined as the unconditional probability of a stationary stochastic
process not crossing its mean in period \( t \). A non-parametric estimator of \( \gamma \), denoted by \( \hat{\gamma} \), is also suggested and its theoretical distributional properties investigated.

In particular, it is shown that \( \hat{\gamma} \) is an unbiased estimator of \( \gamma \), when the mean of the time series process is known and a consistent estimator of \( \gamma \) when the mean is unknown. Inference on \( \gamma \) may be conducted resorting to the conventional approach in which a consistent kernel estimator for the asymptotic variance of \( \hat{\gamma} \), \( \sigma_{\hat{\gamma}}^2 \), is used, or following the recent approach suggested in Kiefer and Vogelsang (2002), in which a non-consistent kernel estimator for \( \sigma_{\hat{\gamma}}^2 \) is used to construct a statistic with a non-standard distribution.

The relationship between \( \gamma \) and other measures of persistence with a particular focus on the "sum of the coefficients" in a pure autoregressive process, which we denote by \( \rho \), is also investigated. It is shown that there is a monotonic relationship between \( \rho \) and \( \gamma \) (and some other scalar measures of persistence) when the data are generated by an AR(1) process, but such a monotonic relationship ceases to exist once higher order autoregressive processes are considered.

The finite sample performance of \( \hat{\gamma} \) and \( \hat{\rho} \), the OLS estimator of \( \rho \), is compared using some Monte Carlo experiments. It is seen that \( \hat{\gamma} \), which has the nice property of being immune to potential model misspecifications, is not significantly affected by the presence of outliers in the data, and the coverage ratio of their empirical confidence intervals is on par with the ones obtained for \( \hat{\rho} \), when one uses the method proposed by Kiefer and Vogelsang (2002).

Finally, the use of the new measure of persistence is illustrated by evaluating inflation persistence in the United States and the Euro Area. We find that, conditional on a break in the mean of inflation, the U.S. and the E. A. do not differ significantly, as far as inflation persistence for the period 1984-2008 is concerned. Both countries exhibit low levels of persistence and there is no significant evidence that the degree of persistence has changed over time with the average level of inflation. However, when we look for changes in persistence not related to changes on average inflation, the use of \( \gamma \) allows
us to uncover a significant reduction of inflation persistence in the Euro Area, occurring after 2001/2002, soon after the launch of the euro and the implementation of a common monetary policy by the European Central Bank. This reduction is such that persistence in the Euro Area becomes virtually nil for the period 2002-2008.

The rest of the paper is organized as follows. Section 2 introduces $\gamma$ and its non-parametric estimator $\hat{\gamma}$ and derives the distributional properties of $\hat{\gamma}$. Section 3 discusses the relationship between $\gamma$ and alternative measures of persistence in the context of the AR(1) and AR(2) models. Using Monte Carlo simulations, section 4 investigates the finite sample properties of $\hat{\gamma}$ and $\hat{\rho}$. Section 5 illustrates the use of the new measure of persistence by evaluating the persistence of inflation in the United States and the Euro Area, and section 6 concludes.

2 An alternative measure of persistence

For the purpose of this paper persistence of a stationary time series, $y_t$, is defined as the speed with which $y_t$ converges to its equilibrium or long-run level after a shock\textsuperscript{2}. In the context of the so-called univariate approach, persistence is investigated by looking at the univariate time series representation of the data. For that purpose it is usually assumed that the data are generated by a stationary autoregressive process of order $p$, (AR($p$)), which may be written as

$$y_t = \alpha + \sum_{j=1}^{p} \beta_j y_{t-j} + \varepsilon_t$$

and reparameterised as:

$$\Delta y_t = \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + (\rho - 1)[y_{t-1} - \mu] + \varepsilon_t$$

\textsuperscript{2}This definition is similar to other definitions in the literature under the assumption of stationary processes (see, for instance, Andrews and Chen, 1994, Willis, 2003, or Pivetta and Reis, 2007). For a different definition of persistence, especially suited for I(1) processes see Jaeger and Kunst (1990).
where
\[ \rho = \sum_{j=1}^{p} \beta_j; \quad \delta_j = -\sum_{i=1}^{p} \beta_i \]
and \( \mu = \alpha/(1 - \rho) \) is the unconditional mean of the series. In the context of model (1), \( y_t \) is said to be (highly) persistent if, following a shock to the disturbance term \( \varepsilon_t \), \( y_t \) converges slowly to its mean (which may be seen as representing the equilibrium level of \( y_t \)). Thus, in the context of this parametric representation of \( y_t \), the concept of persistence appears as intimately linked to the impulse response function (IRF) of the AR(p) process.

It has been argued (see, Andrews and Chen, 1994) that the cumulative impulse response (CIR) is generally a good way of summarizing the information contained in the impulse response function (IRF) and thus, a good scalar measure of persistence. In a simple AR(p) process we have \( CIR = 1/(1 - \rho) \) where \( \rho \) is the “sum of the autoregressive coefficients”, as defined in (3). This monotonic relationship between the CIR and \( \rho \) explains why \( \rho \) may be used a measure of persistence. Thus, using the CIR or simply \( \rho \) as a measure of persistence amounts at measuring persistence as the sum of the disequilibria (deviations from equilibrium) generated during the whole convergence period. The larger is \( \rho \), the larger the cumulative impact of the shock will be.

### 2.1 A new measure of persistence

It is well-known that stationary processes display mean-reversion. In equation (2) the presence of mean reversion is reflected in the term \( (\rho - 1)[y_{t-1} - \mu] \). This implies that if in period \((t-1)\) the series \( y \) is above (below) the mean, the deviation \([y_{t-1} - \mu]\) will contribute as a driving force to a negative (positive) change of the series in the following period, through the coefficient \((\rho - 1)\), thus bringing it closer to the mean. Of course, everything else constant, mean reversion will increase when the coefficient \((\rho - 1)\) increases (in absolute terms). Given that we can measure persistence by \( \rho \) and
mean reversion by \((\rho - 1)\), we conclude that mean reversion and persistence are inversely related: high persistence implies low mean reversion and vice-versa.

We may now introduce a new measure of persistence, which we denote by \(\gamma\), and define as the *unconditional probability of a stationary stochastic process \(y_t\) not crossing its mean in period \(t\). By noticing that \(y_t\) does not cross the mean, \(\mu\), in period \(t\) if and only if \((y_t - \mu)(y_{t-1} - \mu) > 0\), we may formally define \(\gamma\) as

\[
\gamma = P \left\{ [(y_t - \mu) > 0 \land (y_{t-1} - \mu) > 0] \lor [(y_t - \mu) < 0 \land (y_{t-1} - \mu) < 0] \right\}
\]  

Intuitively, the use of \(\gamma\) as a measure of persistence may be justified as a simple implication following directly from the very definition of persistence. If a persistent series is the one which converges slowly to its equilibrium level (the mean) after a shock, then such a series, by definition, must exhibit a low level of mean reversion, i.e., must cross its mean infrequently. Similarly, a non-persistent series must revert to its mean very frequently. And \(\gamma\) simply measures how infrequently a given stationary process crosses its mean.

In contrast to \(\rho\), which requires the data generating process (DGP) to follow a pure autoregressive process, \(\gamma\) is defined independently of the specific underlying DGP, provided stationarity is assumed. In this sense \(\gamma\) as a measure of persistence is broader in scope than \(\rho\).

From definition (4) it is clear that in order to compute the value of \(\gamma\) we need to know the joint probability density function (p.d.f.) of \(y_t\) and \(y_{t-1}\). Obtaining \(\gamma\) is straightforward when \(y_t\) follows a stationary ARMA(p,q) process with normal innovations. In fact, denoting the joint cumulative distribution function of \(y_t\) and \(y_{t-1}\) by \(F(y_t^*, y_{t-1}^*) = P(y_t \leq y_t^* \land y_{t-1} \leq y_{t-1}^*)\) and noticing that the bivariate normal distribution is symmetric relative to any line that crosses the mean and divides the Cartesian
plane evenly, we get $\gamma = 2 \times F(0; 0)^3$. In the next section we shall present the values of $\gamma$ for alternative stationary AR(1) and AR(2) processes.

2.2 A nonparametric estimator for $\gamma$

In empirical applications the DGP is not known so $\gamma$ is also not known. We suggest estimating $\gamma$ by

$$\hat{\gamma} = 1 - \frac{n}{T}$$

(5)

where $n$ stands for the number of times $y_t$ crosses the mean during a time interval with $T + 1$ observations$^4$.

Since $\hat{\gamma}$ is computed by counting the number of mean crossings irrespective of the true underlying DGP, $\gamma$ may be used as a measure of persistence without requiring the researcher to specify and estimate a model for the $y_t$ series. In empirical applications this may imply an important advantage of $\gamma$ over parametric scalar measures of persistence such as $\rho$ or the half-life. In fact, by construction, $\hat{\gamma}$ is immune to potential model misspecifications and, given that it is a non-parametric statistic it can be expected to be robust against outliers in the data. We shall investigate the robustness issue below in section 4.

Note that $\gamma$, by definition, and $\hat{\gamma}$, by construction, are always between zero and one. In the next section it will be shown that for a symmetric zero mean white noise process (zero persistence process) we have $\gamma = 0.5$, so that values of $\hat{\gamma}$ close to 0.5 signal the absence of any significant persistence while figures significantly above 0.5 signal

$^3$Obtaining $\gamma$ may be more difficult under a different assumption for the distribution of the innovations. But, from an empirical point of view, this is not a serious problem. Below we suggest an estimator of $\gamma$ which is an unbiased estimator when the mean of the process is known, so that a fairly good approximation to $\gamma$ may be obtained using Monte Carlo simulations for any assumed distribution of the innovations.

$^4$In theory an estimate of $\gamma$ could also be computed from an estimated Arma model following the steps presented above for $\gamma$. But this is would not be a practical procedure as it would require the specification and estimation of a model. Moreover it would make the properties of the estimator of $\gamma$ dependent upon the properties of the estimator of the model.
significant persistence. On the other hand, figures below 0.5 signal negative long-run autocorrelation.

Deriving the asymptotic distribution of $\hat{\gamma}$ is straightforward. Let us assume that we have a sample with $T+1$ observations, denoted $y_0, y_1, ..., y_T$, generated by a stationary and ergodic process with known mean, $\mu$, and define $x_t$ $(t=1, 2, ..., T)$ such that $x_t$ equals 1 if $y_t$ crosses the mean in period $t$ and is zero otherwise. Given that $\hat{\gamma} = 1 - \bar{x}$ where $\bar{x}$ is the sample mean of $x_t$ it follows that all the results available in the literature concerning consistence and asymptotic distribution of the sample mean of $x_t$ apply directly to the $\hat{\gamma}$ statistic.

We know that if $y_t$ is a stationary and ergodic process $x_t$ is also a stationary and ergodic process, because it is a measurable function of the current and past values of $y_t$. Thus it follows from the law of large numbers that $\hat{\gamma}$ is a consistent estimator of $\gamma^5$. And, in the important special case when $x_t$ is a covariance stationary process with known mean, it turns out that $\hat{\gamma}$ is an unbiased estimator of $\gamma^6$.

When the mean of $y_t$ is unknown the true $x_t$ process is also unknown. What we know is $x_t^*$, which differs from $x_t$ to the extent that the use of $\bar{y}$ instead of $\mu$ may imply some additional mean crossings in $x_t^*$ which are not present in $x_t$. However when $T \to \infty$, $\bar{y}_T \to \mu$ so that $x_t^* \to x_t$ and consistency of $\hat{\gamma}$ follows.

Under the assumption that $x_t$ (or $x_t^*$) is a covariance stationary process it follows by the central limit theorem that $\hat{\gamma}$ is asymptotically normal distributed$^7$, i.e.,

$$\frac{\sqrt{T}(\hat{\gamma} - \gamma)}{\sqrt{\sigma_{\hat{\gamma}}^2}} \overset{d}{\to} N[0, 1] \quad (6)$$

---

$^5$See, for instance, White (1984), Theorems 3.34 and 3.35.
$^6$See, for instance, Hamilton (1994), Chap.7, section 7.2.
where \( \sigma^2 \) the asymptotic variance of \( \sqrt{T} (\hat{\gamma} - \gamma) \) is given by

\[
\sigma^2 = \lim_{T \to \infty} T \cdot E(\hat{\gamma} - \gamma)^2 = \sum_{j=-\infty}^{\infty} r_j = r_0 + 2 \sum_{j=1}^{\infty} r_j
\]  

with \( r_j = \text{cov}(x_t, x_{t-j}) \).

In the special case in which \( y_t \) follows a symmetric zero mean white noise process (zero persistence) (6) reduces to:

\[
\frac{\sqrt{T} (\hat{\gamma} - 0.5)}{\sqrt{0.5}} \overset{d}{\to} N(0, 1)
\]  

which allows carrying out some simple tests on the statistical significance of the estimated persistence (i.e., \( \gamma = 0.5 \)).

In order to implement (6) in practice a consistent estimator for \( \sigma^2 \) can be obtained following a nonparametric estimator of the form

\[
s^2(\gamma)(m) = \frac{\sigma^2}{T} \sum_{j=1}^{\gamma} w_m(j) \hat{r}_j
\]  

where \( \hat{r}_j = T^{-1} \sum_{t=j+1}^{T} (x_t - \bar{x})(x_{t-j} - \bar{x}) \) is the estimate for the j-th order autocovariance of \( x_t \), and \( w_m(.) \) is a lag-window or kernel function depending on a bandwidth parameter \( m \). Examples of conventional kernels often used in the literature are the Bartlett, Parzen and the Quadratic Spectral kernel (for a discussion see, for instance, Andrews, 1991, or den Haan and Levin, 1997).

An alternative approach to get an estimate of \( \sigma^2 \) is the use of bootstrap time series techniques, especially designed for estimating the variance of the sample mean of a stationary process (see, Künsch, 1989, Liu and Singh 1992, Politis and Romano, 1992, 1994 and Paparoditis and Politis 2001).

Recently, Kiefer and Vogelsang (2002) [KV (2002) hereafter] proposed an alternative approach to make inference on the parameters of a model, when the residuals are serially
correlated. Their results also allow to conduct inference on $\gamma$ using the estimator \( \hat{\gamma} \). The method is easy to implement as it basically requires to estimate $\sigma^2_{\hat{\gamma}}$ by some kernel method using the bandwidth parameter equal to sample size in (9). Although this estimator is not consistent for $\sigma^2_{\hat{\gamma}}$, the authors show that the t-ratio

$$t^* = \frac{\sqrt{T(\hat{\gamma} - \gamma)}}{\sqrt{s^2_{\hat{\gamma}}(T)}}$$

(10)

where $s^2_{\hat{\gamma}}(T)$ corresponds to (9) with a bandwidth equal to $T$, converges to a non-standard distribution that does not depend on nuisance parameters. Below resorting to Monte Carlo techniques, we shall investigate the finite sample properties of $\hat{\gamma}$ including the performance of (6) and (10) when the Bartlett window is used.

### 2.3 Testing for changes in persistence using $\gamma$

By noticing that $\hat{\gamma}$ can be obtained as the sample mean of a stationary series, $x_t$, testing for changes in $\gamma$ using $\hat{\gamma}$ is equivalent to test for changes in the mean of $x_t$ using the sample mean as its estimator. This implies that all the tests available in the literature that allow testing for breaks in the mean of stationary processes may be used to test for changes in $\gamma$. Examples include Andrews (1993), Bai (1994, 1997), Bai and Perron (1998) and Altissimo and Corradi (2003) which allow testing for breaks at unknown points in the sample. If one whishes to test for a single change in persistence in a specific (exogenously determined) date we may directly use (6) or (10). Suppose we are investigating persistence for the period $t = 1, 2, \ldots, T$ and we want to test whether there is a change in persistence occurring in period $t = s$, such that persistence for the sub-period $t = 1, 2, \ldots, s$ differs from persistence for the sub-period $t = s + 1, \ldots, T$. By noting that $\hat{\gamma}$ can be computed by regressing $x_t$ on a constant to test for a change in
persistence it suffices to estimate the model

\[ x_t = \alpha_1 + \alpha_2 d_t + u_t \]  

where \( d_t \) is a dummy variable which is zero until the date of the break \((t \leq s)\) and equals 1 thereafter \((t > s)\). In (11) we have \( \alpha_1 = 1 - \gamma_1 \) and \( \alpha_2 = \gamma_1 - \gamma_2 \) where \( \gamma_1 \) and \( \gamma_2 \) are the measures of persistence for the first and second sub-period, respectively. Thus, testing whether persistence has changed amounts to testing whether \( \alpha_2 \) is significantly different from zero in (11). This test can be done using (6) or (10), properly adapted to this specific situation.

2.4 The literature on the frequency of mean crossings

The statistic \( \hat{\gamma} \) has a strong bearing with some other statistics suggested in the literature in areas involving level-crossing problems, runs tests and unit root tests. The so-called zero-crossing problem, which involves determining the distribution of the length of time between the zeros of a zero-mean stationary random process and computing parameters such as the expected number of zero crossings per unit of time and its variance, has long ago been identified as an important issue in Mathematics and Statistics. Examples of significant contributions that address this problem are Rice (1945), Helstrom (1957), Longuet-Higgins (1962), Itô (1964), Ylvisaker (1965), Wong (1966), Blake and Lindsay (1973), Longuet-Higgins (1962), Itô (1964), Ylvisaker (1965), Wong (1966), Blake and Lindsay (1973), Marcus (1977) and Abrahams (1982, 1986). In a discrete context, the behaviour of the zero crossing rate, which in a zero mean stationary process corresponds exactly to \( 1 - \hat{\gamma} \), has also been investigated in order to establish important properties of the underlying processes. See, for instance, Kedem and Slud (1982), Kedem (1986), He and Kedem (1990) and Cheng et al. (1997).

The estimator \( \hat{\gamma} \) is also closely related to the runs test, which has been used in the literature as a way of testing for the absence of serial correlation. The runs test traces at
least to David (1947) (see also Goodman, 1958, or David and Barten, 1962), a run being an interval in which you are always on one side of the mean, the median or any arbitrary constant. If we take the total number of runs defined with reference to the mean it is immediate to realize that the number of mean crossings equals the total number of runs less one, i.e., \( n = R - 1 \), where \( R \) stands for the total number of runs. Granger (1963) discusses the use of the statistic \( S = (2R/T) - 1 \) suggested in David (1947), as a way of testing for the absence of serial correlation. Given that \( n = R - 1 \), we have \( S = 1 - 2(T-1)\tilde{\gamma}/T \) so that for large \( T \) we get \( S \approx 1 - 2\tilde{\gamma} \). Under the assumption that the distribution of the process is symmetric, the David runs test is such that for sufficiently large \( T \) the distribution of \( R/T \approx 1 - \tilde{\gamma} \) can be well approximated by \( N(1/2, 1/2\sqrt{T}) \) so that we get result (8) above. Thus, using the runs test \( R/T \) is equivalent to using \( \tilde{\gamma} \) to test for the absence of serial correlation under the assumption of a symmetric white noise process.

The use of the frequency of mean crossings to test for a unit root in a discrete time series has also been suggested in the literature (see Burridge and Guerre, 1996 and García and Sansó, 2006). Under the assumption that the series, \( y_t \), is generated by a pure random walk process with no drift \( (y_t = y_{t-1} + \varepsilon_t) \) and initial value equal to zero \( (y_0 = 0) \) Burridge and Guerre (1996) demonstrated that the statistic

\[
K_T^*(0) = \frac{\sqrt{\sum \frac{(\Delta y_t)^2}{T}}}{\sqrt{T}} \cdot \frac{n}{\sqrt{T}} = \omega \cdot n \sqrt{T}
\]

is such that \( K_T^*(0) \overset{d}{\rightarrow} |Z| \), where \( Z \) is a standard normal \( N(0,1) \) and \( n \) is the number of sign changes of \( y_t \) during a time interval with \( T \) observations. It is straightforward to show that \( K_T^*(0) \) can be rewritten in terms of the \( \hat{\gamma} \) statistic as \( K_T^*(0) = \omega \frac{(T+1)}{\sqrt{T}}(1 - \hat{\gamma}) \). Once \( \hat{\gamma} \) converges to 1 as \( \rho \) goes to 1, this statistic can be used to test the null of \( \gamma = 1 \). Notice, however, that (12) is valid only under the assumption that the data are generated by a random walk without drift with \( y_0 = 0 \) and it has lower power than the conven-
tional Dickey-Fuller test (see Burridge and Guerre, 1996). García and Sansó (2006) who generalised the Burridge and Guerre statistic for processes with a deterministic trend and more general innovations, also concluded that their generalization has lower power than the DF$^{GLS}$ test proposed by Elliot et al. (1996).

3 The relationship between $\gamma$ and alternative measures of persistence

In this section we compare $\gamma$ with alternative measures of persistence, in the context of different stationary AR(p) models. Column (2) of Table 1 presents the values of $\gamma$ corresponding to the value of $\rho$ in column (1) assuming that the data are generated by the AR(1) process $y_t = \rho y_{t-1} + \varepsilon_t$ with normal innovations, $\varepsilon_t$. For instance, for the AR(1) process with $\rho = 0$ (white noise process) $\gamma$ is 0.50 while, for the AR(1) process with $\rho = 0.60$ the corresponding $\gamma$ is 0.705. Table 1 shows that, in the context of the AR(1) process, there is a one-to-one correspondence between these two measures of persistence.

Table 1 also reports the half-life, $h$, defined as the number of periods for which the effect of a unit shock remains above 0.5, as well as, $m_{50}, m_{95}$ and $m_{99}$, which denote the number of periods required for the accumulated effect of a unit shock to be equal to 50%, 90% and 99% of the total effect, respectively. We see these measures of persistence as useful tools to evaluate how fast the series approaches the equilibrium following a shock$^8$.

Looking at columns (3) and (4) we realise that $h$ and $m_{50}$ assume exactly the same values for the different values of $\rho$, as one would expect under an AR(1) process. As to the remaining measures of persistence $m_{95}$ and $m_{99}$, they exhibit a monotonic relationship

---

8We compute these four measures of persistence directly from the IRF. Moreover, in contrast to some literature that computes the half-life on a continuous basis, in Tables 1 and 2 we assume a discrete range of variation for this measure of persistence (as well as for $m_{50}, m_{95}$ and $m_{99}$).
with $\rho$ and $\gamma$ (specially so for $m_{99}$). Thus, in the context of the AR(1) model, the different measures of persistence, especially $\rho$, $\gamma$ and $m_{99}$ appear as giving the same message about the degree of persistence. This result stems from the fact that in the AR(1) model the speed of convergence of the IRF to the equilibrium is constant throughout the adjustment period.

In order to investigate whether this relationship among the different measures of persistence carries over to higher order autoregressive models we now consider the AR(2) process, $y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \epsilon_t$. Given the wide range of possibilities for the two autoregressive coefficients we set their sum equal to 0.80 (i.e., $\rho = \rho_1 + \rho_2 = 0.80$). Thus models, (1) to (9) in Table 2 correspond to different combinations of the two AR coefficients that meet the conditions: i) $\rho_1 \geq 0$, ii) $\rho_1 + \rho_2 = 0.80$ and iii) the model is stationary. In addition to the measures of persistence considered in Table 1, Table 2 also includes the “largest autoregressive root”, $lar$, as an additional measure of persistence.\(^9\)

In Table 2 models are listed according to the values of $\rho_1$ in descending order. If we stick to $\rho$ as the single measure of persistence all the 9 models would be seen as equally persistent, with $\rho = 0.80$. However, as we move from model (1) to model (9) we realise that the value of $\gamma$ and of the half-life, $h$, decrease monotonically while, in strong contrast, the values of $lar$, $m_{50}$, $m_{95}$ and $m_{99}$ increase monotonically. Thus, if instead we stick to one of these six alternative measures the 9 models would appear as essentially different in terms of persistence. In other words, the 6 alternative measures of persistence behave quite independently from the value of $\rho$.

This is a very important result as it shows that the conclusion above for the AR(1) model that the alternative measures of persistence basically convey the same message, does not carry out to more general autoregressive processes. The diverging results for the different measures of persistence in the AR(2) process stem from the fact that, in contrast to the simple AR(1) model, the speed of convergence of the IRF throughout

the adjustment period is not constant for higher AR(p) models (the shape of the IRF varies with the specific combination of the autoregressive coefficients).

The idea that the scalar measures of persistence could in some specific situations lead to very different conclusions about the degree of persistence is, of course, not new\textsuperscript{10}. What seems to be new (at least for the authors) is the extension of the problem. More than saying that in some special cases the scalar measures of persistence are not good summary measures of the information contained in the IRF it seems more appropriate to state that with the exception of the very special case of the AR(1) model, the scalar measures of persistence for the general autoregressive model may be very misleading, either because they may suggest the presence of a strong degree of persistence when it is absent or the absence of significant persistence when it is present.

Let us take a further look into this issue. Graph No.1 displays a realization of models (1), (5) and (9) with 80 observations\textsuperscript{11}. The three models despite having the same $\rho = 0.80$, display quite different values for the 6 measures of persistence recorded in Table 2. In particular, we see that average mean reversion in model (5) and model (9) is clearly higher than in model (1) suggesting that persistence, as measured by the lack of mean reversion, is lower in those two models. Moreover, we also see that mean reversion in model (9) is higher than in model (5). From Table 2 we can see that $\gamma$, starting with model (1), decreases monotonically from a value as high as 0.853 (signalling a very persistent process) to a figure as low as 0.50, which signals a model with zero persistence.

We expect $\gamma$ to be equal to 0.50 when a white noise process generates the data, but by simply eyeballing the series we see that model (9) in Graph No.1 does not behave

\textsuperscript{10}Andrews and Chen (1994) discuss several situations in which the CIR and thus also $\rho$ might not be sufficient to fully capture the existence of different shapes in the impulse response function. In turn, Pivetta and Reis (2007) list the main limitations of the “half-life”. The use of the largest autoregressive root as a measure of persistence is criticised both in Andrews and Chen (1994) and in Pivetta and Reis (2007).

\textsuperscript{11}The three series in Graph No.1 were generated using the same series of residuals generated from the N(0,1) distribution.
like a white noise. Rather it seems to display a kind of cyclical behaviour, which makes
the process to cross the mean at irregular intervals, but such that on average it crosses
the mean as often as if it were a symmetric white noise process. This means that \( \gamma \) does
not distinguish between a process with a low \( \rho \) (close to a white noise behaviour) and a
process with a cyclical pattern such that, on average, during a given time interval the
two processes cross the mean the same number of times. In other words \( \gamma \), in contrast
to \( \rho \), does not see the cyclical pattern of the process as relevant persistence\(^{12}\).

Given the evidence in Table 2 that shows that for higher order processes the different
scalar measures of persistence may deliver conflicting views on the degree of persistence,
it seems wise, in empirical applications, not to rely on a single measure of persistence.
In this regard, using \( \gamma \) and \( \rho \) as companion measures of persistence seems to be a good
strategy that may allow uncovering important features of persistence, that would not be
identifiable if we stuck to a single measure. In fact, as we have just seen, a high value of
\( \rho \) accompanied by a low value of \( \gamma \) might be a sign of a cyclical pattern in the DGP. We
also shall see below that an estimate of \( \rho \) clearly below to what could be expected given
the value of \( \hat{\gamma} \) might be a signal of significant downward biases in \( \hat{\rho} \) stemming from the
presence of outliers in the data.

4 Some Monte Carlo evidence on the finite sample
properties of \( \hat{\gamma} \) and \( \hat{\rho} \).

In this section we use some Monte Carlo experiments in order to investigate the prop-
erties of \( \hat{\gamma} \) and \( \hat{\rho} \) regarding i) unbiasedness, ii) robustness to outliers and iii) coverage

\(^{12}\)Notice that a similar situation occurs with \( \rho \). In particular, \( \rho \) does not allow us to discriminate
between a white noise process and any stationary process for which \( \rho = \sum \rho_i = 0 \). For example, the
IRF for the process \( y_t = 0.70y_{t-1} - 0.70y_{t-2} + \varepsilon_t \) differs from that of a white noise process, but \( \rho \), as
a measure of persistence, is unable to tell the difference. In this case, \( \gamma \) would allow us to discriminate
the two processes as we have \( \gamma = 0.50 \) for the white noise and \( \gamma = 0.635 \) for the AR(2) process.
ratio of empirical confidence intervals. These properties are investigated in the context of the AR(1) and AR(2) processes considered in section 3.

From the discussion in section 2 we may expect the information about the mean of the process to be statistically relevant for persistence evaluation. For this reason, below we distinguish the situation in which the mean is known from the situation in which the mean is unknown\textsuperscript{13}.

4.1 Unbiasedness

Let us start by assuming that the true mean of the process is known and define an experiment that constructs the data to follow an AR(1) process (with no intercept) given by $y_t = \rho y_{t-1} + \varepsilon_t$, for $\rho$ ranging between 0 and 1 and where the errors are serially uncorrelated standard normal variables. Samples of size $T = 50, 75, 100, 150, 250, 500$ and 1000 are used in the experiments\textsuperscript{14}.

The output of the experiment for $T=100$ is displayed in Table 3\textsuperscript{15}. For values of $\rho$ ranging from 0 to 0.95, column (2) reports the values of $\gamma$ (taken from Table 1), column (3) reports the average value for the Monte Carlo OLS estimates of $\rho (\bar{\rho})$ and column (5) the average value of the Monte Carlo estimates of $\gamma (\bar{\gamma})$. From columns (1) and (3) we can see that the OLS estimator of $\rho$ is slightly (mean) downward biased and that the absolute bias increases as $\rho$ increases, as expected (see, Sawa, 1978, Phillips, 1977, Evans and Savin, 1981, Andrews, 1993, Evans and Chen, 1994).

\textsuperscript{13}In the context of an inflation persistence evaluation exercise, assuming that the mean is known may be realistic for those countries for which an inflation targeting monetary policy was implemented and an explicit inflation target was announced. In this case, the true mean of the series can be computed exogenously to realised inflation as the publicly announced inflation target. However, for most countries, the exact (implicit) inflation target used by the central bank when setting monetary policy is unknown. In these cases the mean must be computed from realised inflation.

\textsuperscript{14}As the sampling distribution $\hat{\rho}$ depends on the initial value of the process, $y_0$ (see, for instance, Evans and Savin, 1981) in our simulations we create $T + 100$ observations and discard the first 100 observations in order to remove the effect of initial conditions.

All the experiments are replicated 10,000 times with the data generated by setting $y_{-100} = 0$. The replications were carried out using TSP 5.0.

\textsuperscript{15}The output of the simulations for other values of $T$ is available from the authors upon request.
As regards the \( \hat{\gamma} \) statistic we see that the values of \( \gamma \) and \( \bar{\gamma} \) in columns (2) and (5) are not statistically different i.e., \( \hat{\gamma} \) behaves as an unbiased estimator of \( \gamma \), as expected, given the discussion in section 2\(^{16}\).

Let us now assume that the mean of the process is unknown and thus has to be estimated from the data. The results of this new Monte Carlo experiment are in columns (6)-(9) of Table 3. Looking at column (7) we see that the downward bias of OLS estimator has now significant damaging consequences on the expected estimates of \( \rho \). In fact, the bias has now more than doubled vis-à-vis the situation in column (4), confirming the claim in Sawa (1978) and Andrews (1993) that the bias of \( \hat{\rho} \) is more acute when one assumes that the mean of the process is unknown and has to be estimated from the data using a model with an intercept. In empirical applications this downward bias of \( \hat{\rho} \) will naturally translate into all measures of persistence that are computed using an estimate of \( \rho \) (\( h, m_{50}, m_{95} \) and \( m_{99} \), in Table 1). This is the case for which it might be worth using the “approximated median unbiased estimator” suggested in Andrews and Chen (1994).

As to the \( \hat{\gamma} \) statistic we see that it also appears slightly downward biased as expected given the discussion in section 2. This result is intuitive, as the estimator of the mean (the sample average) is expected to increase mean reversion vis-à-vis the situation with the true mean, and thus reduce the estimated \( \gamma \).

Let us now take a look at the AR(2) process. In our Monte Carlo experiment we consider as our DGP’s the same 9 models of section 4 for which \( \rho = \rho_1 + \rho_2 = 0.80 \). The output of the experiment is in Table 4. To facilitate comparisons, column (6) reports the values of \( \hat{\gamma} \) taken from Table 2.

\(^{16}\)Note that the sampling variability allows \( \bar{\gamma} \) (and \( \bar{\rho} \)) to change for a fixed \( T \). As the standard error of \( \bar{\gamma} \) may be approximated by \( \sqrt{0.5/T} / \sqrt{10000} \) and \( \bar{\gamma} \) has a Normal distribution, a 95% confidence interval for \( \bar{\gamma} \) is given by \( \bar{\gamma} \pm 1.96 \times (0.5/\sqrt{T}) / \sqrt{10000} \), which reduces to \( \bar{\gamma} \pm 0.001 \) for \( T = 100 \). From the output of the simulations we see that the values of \( \bar{\gamma} \) do not differ from \( \gamma \) by more than 0.001, so that the difference can be attributed to sampling variability.
By looking at columns (2) and (3), we see that the biases of $\hat{\rho}$ increase monotonically as we move from model (1) to model (9). This is an interesting result because it shows that for higher order processes the bias of $\hat{\rho}$ depends not only on $\rho$ but also on the specific combination of $\rho_1$ and $\rho_2$. By comparing columns (6) and (7) we see that, as expected, $\hat{\gamma}$ behaves as an unbiased estimator of $\gamma$.

If we assume that the mean of the process is unknown and estimate a model with an intercept we find that, as expected, the downward biases of the OLS estimator of $\rho$ increase significantly (columns (4) and (5)). As regards the $\hat{\gamma}$ statistic we find that, similarly to the AR(1) case, a small downward bias emerges (columns (8) and (9)) but it is always smaller than the bias displayed by $\hat{\rho}$.

Thus, from the preceding analysis, we conclude that in the process of persistence evaluation it may be worth distinguishing between two different possibilities. When the mean is known we may expect to be able to estimate persistence with no (expected) bias if $\hat{\gamma}$ is used or, with a small downward bias if $\hat{\rho}$ is used (in this latter case, assuming also that the order of the process is known). However, when the mean is estimated from the data, which corresponds to the common practice in the literature, we may expect such a fact to introduce an additional downward bias into the conventional measures of persistence. This bias might be particularly significant if OLS estimators are used to get an estimate of $\rho$. The $\hat{\gamma}$ statistic is very much less affected in such a situation.

4.2 Robustness to outliers

In order to evaluate the robustness of $\hat{\gamma}$ and $\hat{\rho}$ to the presence of outliers in the data we define the DGP as corresponding to the AR(1) model without an intercept used in sub-section 4.1, with the addition of 5% of observations drawn from the $N(0, 5^2)$ distribution\textsuperscript{17}. Table 5 reports the results for $T = 100$. Column (4) presents the estimated

\textsuperscript{17}The type of outliers we consider are those that correspond to shocks that affect observations in isolation due to some non-repetitive events, which may occur as a result of measurement errors or special events (changes in VAT rates or union strikes, for instance). This type of outliers, usually
bias for \( \hat{\rho} \) measured as a percentage deviation from the estimated values obtained in the absence of outliers (see column (3) in Table 3). This way we are measuring only the bias due to outliers.

The first important comment is that the presence of additive outliers in the data has a devastating effect on the OLS estimators. For instance, when \( \rho = 0.60 \) the average estimated \( \hat{\rho} \) (\( \bar{\rho} \) in column (3)) is as low as 0.359 (it is equal to 0.589 when no outliers are present) which corresponds to a downward bias of 39.07%. The effect of outliers decreases as \( \rho \) increases, but even for values of \( \rho \) as large as 0.80 the average \( \hat{\rho} \) is only 0.56.

As regards the \( \hat{\gamma} \) statistic, the estimated bias (measured as a percentage deviation from the estimated values obtained in the absence of outliers in column (5) of Table 3) is reported in column (7). We can see that there is some downward bias as expected (given that some outliers will imply additional crossings of the mean), but it is quite small. For instance, for the model with \( \rho = 0.60 \) the average \( \hat{\gamma} \) is now 0.689 while it was 0.705 when no outliers were present. In general the bias of \( \hat{\gamma} \) due to outliers increases as \( \rho \) increases but it always remains very small.

If instead we take a look at the estimated standard errors of both \( \hat{\rho} \) and \( \hat{\gamma} \) (not shown in the Tables) and compare them to the corresponding standard errors obtained in the absence of outliers, we conclude that the implications are much stronger for the standard errors of \( \hat{\rho} \). In fact, while the standard errors of \( \hat{\gamma} \) show a small increment the standard errors of \( \hat{\rho} \) for higher values of \( \rho \) more than doubled. The implications for the standard deviations would naturally be reflected, for instance, in the properties of the interquartile range of each estimator. From column (5) we see that, with the exception of the model with \( \rho = 0.00 \), the interquartile range of the OLS estimator does not include the true \( \rho \) (nor the estimated \( \rho \) when no outliers are present). In contrast, the interquartile range referred to in the literature as additive outliers, is known to have strong impacts on the parameters of estimated models (see, for instance, Lucas, 1995).
for \( \hat{\gamma} \) (column (8)) always includes the true \( \gamma \) (or the estimated \( \gamma \) obtained in Table 3 when no outliers are present).

Results in Table 5 are of course specific to the particular way we generate the data, and less extreme outliers are expected to have less damaging consequences for the estimators. However, the exercise carried out shows that in general we can expect \( \hat{\gamma} \) to be more robust to the presence of additive outliers in the data than \( \hat{\rho} \).

### 4.3 Coverage ratio of empirical confidence intervals

In order to evaluate the consequences for inference stemming from the use of empirical estimates of the asymptotic variance of \( \hat{\gamma} \) in finite samples, now we compute the coverage ratio of 95\% confidence intervals for \( \rho \) and \( \gamma \) using the same AR(1) and AR(2) models, as before.

We restrict our simulations to the Bartlett kernel-based estimator as it is probably the most widely used kernel in so-called heteroskedasticity and autocorrelation consistent estimation and also the one that underlies some well-known nonparametric unit root and stationary tests (see Phillips and Perron, 1988, and Kwiatkowski et al., 1992). Thus, to compute the long run variance of \( \hat{\gamma} \) we use the estimator

\[
s^2_{\hat{\gamma}} = T^{-1} \sum_{t=1}^{T} (x_t - \bar{x})^2 + 2T^{-1} \sum_{j=1}^{m_k} \left( 1 - \frac{j}{m + 1} \right) \sum_{t=j+1}^{T} (x_t - \bar{x})(x_{t-j} - \bar{x})
\]

for alternative number of lags, \( m_k \), determined using three different procedures. The first approach, followed for instance in Schwert (1989), defines \( m_k \) as a function of the number of observations such that \( m_k = \text{int} \left[ k(T/100)^{1/4} \right] \), for alternative values of \( k \).\(^{18}\) The second approach uses the automatic data-based procedure suggested in Andrews (1991), using the AR(1) plug-in method. Finally, the third approach follows the methodology suggested in KV (2002), in which the bandwidth parameter is equal to

\(^{18}\)Below we present the results for \( k = 8 \). We carried out simulations also for \( k = 4 \) and \( k = 12 \), but the qualitative conclusions do not change.
the sample size. The two first approaches are examples of the conventional procedure that involves the use of a consistent estimator for the asymptotic variance of \( \hat{\gamma} \), while the KV (2002) approach uses an estimator for the variance of \( \hat{\gamma} \) which is not consistent.

The results of the Monte Carlo simulations for the AR(1) and AR(2) models with \( T = 100, T = 250 \) and \( k = 8 \), are in Tables 6 and 7. The confidence intervals for \( \gamma \), in the first two approaches, and for \( \rho \) are constructed using the Normal distribution to approximate the true finite sample distribution of \( \hat{\gamma} \) and \( \hat{\rho} \). In the third approach the confidence intervals for \( \gamma \) are constructed using the critical values supplied in KV (2002)\(^{19}\).

The first important point to note regarding the use of \( \gamma \) is that the first two conventional approaches deliver quite different results vis-à-vis the third approach. On the one hand, the exercise shows, for the type of models estimated, that the Bartlett estimator with a predetermined bandwidth parameter, usually underestimates the asymptotic variance of \( \hat{\gamma} \) so that the effective coverage ratio is below its nominal level\(^{20}\). Strangely enough the Andrews approach despite defining the bandwidth parameter as a function of the autoregressive structure of the data, does not significantly improve on the previous approach, notably so for higher values of \( \rho \) in the AR(1) case. Not surprisingly, given the simulations performed by the authors, the estimator suggested in KV (2002) performs significantly better than the other two, delivering confidence intervals for \( \gamma \) whose effective coverage ratio is closer to its nominal level. This of course also means that tests on \( \gamma \) based on the two conventional approaches tend to be oversized, especially so for larger values of \( \rho \) and \( \gamma \), but tests on \( \gamma \) based on the KV (2002) approach have good size properties.

If we consider the relative performance of \( \hat{\rho} \) and \( \hat{\gamma} \) we see that \( \hat{\rho} \) performs better than \( \hat{\gamma} \) when inference on \( \hat{\gamma} \) is made using the two conventional methods, but not when

---

\(^{19}\)In the present case 4.771 is used as the critical value to construct the confidence intervals.

\(^{20}\)This outcome accords with the available Monte Carlo evidence for stationarity tests that rely on kernel estimators (see, for instance, Kwiatkowski et al. 1992, Lee, 1996, Caner and Kilian, 2001).
inference on $\hat{\gamma}$ is made using the KV(2002) approach. In fact, in this latter case, if anything, the coverage ratio for $\gamma$ emerges as slightly better than the coverage ratio for $\rho$. Notice also that the performance of $\hat{\rho}$ and $\hat{\gamma}$ decreases not only as $\rho$ and $\gamma$ increase in the AR(1) case, as expected, but also as we move from model (1) to model (9) in the AR(2) case, despite $\rho$ being held constant. Besides the increased biases in the estimators documented in sub-section 4.1, such an outcome must also be reflecting increased difficulties in the estimation of the variance of $\hat{\rho}$ and $\hat{\gamma}$. As expected, when $T$ increases the difference in the performance of the different approaches is reduced.

An interesting final question regards the power of the tests involving $\hat{\gamma}$. According to simulations in KV(2002) the Bartlett kernel estimator, among the common choices of kernels, produces the highest power function, when the bandwidth is equal to the sample size. However, the power of the tests based on the KV (2002) approach seems to be less than that which can be attained using conventional procedures involving consistent estimators for the asymptotic variance of $\hat{\gamma}$. In fact, according to the simulations in Kiefer and Vogelsang (2005) for kernel and bandwidth choice there is a trade-off between size distortions and power. Smaller bandwidths lead to tests with higher power but greater size distortions while large bandwidths lead to tests with lower power but less size distortions.

5 Persistence in the United States and the Euro Area.

Inflation persistence has been an intensely investigated topic in macroeconomics, over the last decade. Issues such as whether inflation is persistent or not, whether it has changed

\footnote{Notice that the exercise performed here is very benevolent for $\rho$ as it is conducted by estimating the correct model. In empirical applications the relative performance of $\hat{\rho}$ in terms of its coverage ratio may be expected to be significantly weakened because the true model is unknown (and this would generally imply additional biases for $\hat{\rho}$) and the data may contain some outliers (which, as we have seen, have a stronger negative impact on the performance of $\hat{\rho}$).}
over time, whether it is structural or may vary according to the specific monetary policy regime, are examples of relevant questions that have been addressed in the literature. As far as the United States (U.S.) and the Euro Area (E.A.) are concerned, there is now a vast number of contributions aiming at quantifying the degree of inflation persistence and investigating whether it has changed over time. Burdekin and Siklos (1999), Bleaney (2001), Stock (2001), Willis (2003), Levin and Piger (2004), Pivetta and Reis (2007) and Cogley and Sargent (2001, 2007) for the U.S., Gadzinski and Orlandi (2004), Levin and Piger (2004), Corvoisier and Mojon (2005), O’Reilly and Whelan (2005), Altissimo et al. (2006), Angeloni et al. (2006), and Benati (2008) for the E.A., are examples of important contributions.

For the E.A., an important issue is whether the emergence of the Monetary Union and the implementation of a single monetary policy as of 1999 has brought about a significant decline in inflation persistence. From the theoretical front, there are reasons to believe that a decline in persistence could be expected, due to the anchoring of inflation expectations. For instance, according to Altissimo et al. (2006), "by conducting monetary policy such that inflation expectations of economic agents are well anchored, the central bank can ensure that actual inflation does not deviate far too long and in a too persistent fashion from what it has announced as its medium-term objective for inflation". Notwithstanding, from the empirical front there seems to have emerged some conflicting evidence. Angeloni et al. (2006) using $\rho$ and quarterly data for the period 1984-2004 conclude that there is no evidence of a change in persistence around 1999 and that, if anything, inflation persistence may actually have slightly increased. Benati (2008) using $\rho$ and quarterly data for the period 1971-2006 finds strong evidence of a decline in persistence when comparing the sub-period 1971-1998 with the sub-period 1999-2006 (the median unbiased estimate of $\rho$ declines from 1.01 to 0.35). However, the fact that the estimate of $\rho$ for the sub-period 1971-1998 is 1.01 suggests that inflation was not stationary during this length of time. This, of course, casts strong doubts on the
underlying assumption of a single monetary policy regime (i.e., a constant mean) for this
sub-period and thus makes the comparison between the two sub-periods questionable\textsuperscript{22}.

In this section we contribute to this empirical literature by investigating the degree
of inflation persistence in the U.S. and the E.A. using $\gamma$ and $\rho$ as alternative measures
of persistence. Inflation is measured by the first difference of logged GDP deflator for
the period 1984-2008\textsuperscript{23}. The purpose of the analysis is threefold. Firstly, to illustrate
the use of $\gamma$ as a measure of persistence. Secondly, to investigate if persistence varies
with the level of inflation and thirdly, to shed some more light on whether a break in
persistence has occurred in the E.A. with the emergence of the Monetary Union and the
implementation of a single monetary policy.

A major issue when evaluating the degree of inflation persistence regards the way the
mean of inflation is dealt with. Assuming a constant or a time varying mean for inflation
makes all the difference for the estimated persistence. The literature has addressed this
issue by testing for breaks in the mean and computing persistence conditional on such
and Orlandi, 2004, Corvoisier and Mojon, 2005), by dividing the samples according to
historically different monetary policy regimes (Alogoskoufis and Smith, 1991, Burdekin
\textsuperscript{\textsuperscript{\textsuperscript{22}}Assuming that inflation is determined by monetary policy in the long run has the implication
that the stochastic process for inflation must exhibit mean reversion, when the mean is measured by
the central bank inflation target. Thus, the possibility of inflation displaying a unit root in a stable
monetary policy regime has to be excluded, by definition.

Additionally, there is evidence of significant changes in the mean of inflation for most OECD countries
multiple break tests concludes for the existence of three waves of breaks, the first wave occurring in the
late 1960’s or early 1970’s, the second taking place in the first half of the 1980’s and the third wave in
the early nineties. These findings are consistent with the conclusions in Gadzinski and Orlandi (2004)
and Levin and Piger (2004) who investigate a smaller sample (1984 onwards) and also find a wave of
breaks in the mean of inflation in the early nineties. Levin and Piger (2004) notice that the breaks in
the early 1990’s coincide with the spreading of inflation targeting, while Corvoisier and Mojon (2005),
recall that half of the OECD countries, that eventually adopted the euro, have pursued lower levels
of inflation in the nominal convergence process foreseen in the Maastricht treaty signed in 1992.

The existence of a wave of breaks in the mean of inflation in the early 1980’s led many authors to
focus on the most recent period (Gadzinski and Orlandi, 2004, and Levin and Piger, 2004) and explains
why, in this section, we restrict the analysis to the period 1984-2008.

\textsuperscript{23}The data on GDP deflator for the U.S. were downloaded from the Bureau of Economic Analysis
website while that for the E.A. were constructed by updating the data in Fagan et al. (2001) with the
ECB official series from 1995q1 onwards. Both series are seasonally adjusted.
and Siklos, 2001, Benati, 2008) or allowing for a pure time varying mean of inflation as a way of capturing shifts over time in the central bank inflation target (Smets and Wouters, 2003, Cogley and Sargent, 2001, Cogley et al., 2008). Here, we follow the first approach by conducting break tests in the mean of inflation and computing persistence conditional on the estimated means.

To test for a break in the mean we resort both to the Andrews (1993) and the Altissimo and Corradi (2003) tests that allow testing for changes in the mean at unknown points in the sample\textsuperscript{24}. To run the Andrews test we impose 20% symmetric trimming to avoid detecting spurious breaks at the beginning or at the end of the sample. The Andrews test detects a break in 1991q2 for the U.S. and a break in 1993q3 for the E.A., while the Altissimo and Corradi test detects a break in 1991q1 for the U.S. and a break in 1993q2 for the E.A..\textsuperscript{25} Given that the analysis does not change in any significant way, in what follows we stick to the break dates identified by the Andrews tests.

Graph No.2 displays inflation for the U.S. and the E.A. together with their (constant) means for the period 1984q1-2008q4. The upper panels of Graphs No.3 and No.4 display inflation for the U.S. and the E.A. allowing for a break in the mean, at the above identified dates, while the lower panels of these two Graphs depict the corresponding deviations from the mean.

Table 8 displays the main results for the two series\textsuperscript{26}. Column (1) displays the full sample estimates of \( \gamma \) and \( \rho \) under the assumption of a constant mean, while column (2) reports the corresponding estimates under the assumption of a break in the mean. Below \( \hat{\gamma} \) and \( \hat{\alpha}_2 \) are the KV(2002) standard-errors obtained with the Bartlett window.

\textsuperscript{24}We thank Fillipo Altissimo and Valentina Corradi for sharing with us the code of the Altissimo and Corradi (2003) test.
\textsuperscript{25}These two series (for a smaller sample) have been studied in Gadzinski and Orlandi (2004) who found the breaks in the mean to occur in 1991q2 for the U.S. and 1993q2 for the E.A.; 1991q2 is also the date of the break identified in Levin and Piger (2004) for the U.S..
\textsuperscript{26}Estimates of \( \gamma \) were obtained by using equation (5) in Section 2, while estimates of \( \rho \) were obtained by estimating an autoregressive model for the series of the deviations from the mean. For the autoregressive model a general-to-specific approach was followed which delivered a model with 5 lags for the U.S. and with 3 lags for the E.A.
Conditional on a constant mean we get point estimates for persistence that are somewhat higher for the E.A. than the U.S. ($\hat{\gamma}=0.788$ and $\hat{\rho}=0.646$ and $\hat{\gamma}=0.749$ for the U.S.), but the null of equal persistence in the two countries may not be rejected for a 5% test.

If we look at persistence conditional on a break in the mean for the whole sample period, in column (2), we find that estimated persistence is now significantly lower as expected ($\hat{\gamma}=0.677$ and $\hat{\rho}=0.526$ for the E.A., $\hat{\gamma}=0.545$ and $\hat{\rho}=0.666$ for the U.S.) and that once again the null of equal persistence in the two countries may not be rejected for a 5% test. These results suggest that, irrespective of the measure used, persistence of inflation in the E.A. and the U.S. is not high and does not differ significantly between the two countries. Of course, we must be aware that these conclusions are conditional on an estimated mean, which as we have seen in section 4 is likely to introduce some downward bias into the estimators of our measures of persistence.

Column (3) displays the results of a Chow-type test for a change in persistence, occurring at the dates identified for the break in the mean: 1991q2 for the U.S. and 1993q3 for the E.A.. This test allows us to investigate whether persistence differs for high and low levels of inflation. From Table 8 we see that the null of unchanged persistence for the two sub-periods cannot be rejected both for the U.S. and the E.A., irrespective of whether we use $\gamma$ or $\rho$ to measure persistence (the t-ratios of $\hat{\lambda}$ are clearly below the 5% critical value of the normal distribution and the t-ratios of $\hat{\alpha}_2$ are below the 5% critical value reported in Table 1 of KV (2002)). Thus, from this exercise we conclude that

---

27 Tests on $\gamma$ were performed as explained in section 2, by estimating equation (11) and computing the variance of $\hat{\alpha}_2$ using the KV (2002) estimator with the Bartlett window. Tests of a change in $\rho$ were performed by estimating autoregressive models that allow for the possibility of a break in the persistence parameter. In particular, the estimates $\hat{\lambda}$ in Table 8 were obtained from the model

$$z_t = \sum_{j=1}^{p-1} \delta_j \Delta z_{t-j} + \rho_{1} z_{t-1} + \lambda d_t z_{t-1} + \varepsilon_t$$

where $z_t$ is the series of inflation (deviations from the mean) and $d_t$ is a dummy variable which is zero until the date of the break ($t \leq s$) and equals 1 thereafter ($t > s$). More general models were also estimated (for example allowing for the possibility of changes in the $\delta_j$ parameters) but the conclusions do not change.

---

28
for this sample period there is no significant evidence that persistence has changed over time with the average level of inflation.

This result does not preclude the possibility of changes in persistence having occurred in a different part of the sample, not directly related to changes in the average level of inflation. To investigate this possibility we tested for changes in persistence occurring at an unknown point in the sample using again the Andrews and the Altissimo and Corradi tests. Under the assumption of a break in the mean, the two tests still do not find any evidence of a change in persistence for the U.S.. However, the results for the E.A. are quite different according to whether one uses $\rho$ or $\gamma$ as the relevant measure of persistence. The use of $\rho$ does not suggest any change in persistence, but for $\gamma$ both the Andrews and the Altissimo and Corradi tests suggest that a change in persistence has occurred (in 2002q2 according to the Andrews test, in 2001q4 according to the Altissimo and Corradi test).

We can get some intuition for this result by looking at Graph 4. Mean reversion is clearly stronger for the period 2002-2008 than it is in the period before (1984-2001) which is a clear sign that persistence has changed around 2001/2002. In fact, for the period 1984q1-2001q4 we get $\hat{\gamma} = 0.803$ (sd. 0.022) while for the period 2002q1-2008q4 we get $\hat{\gamma} = 0.357$ (sd. 0.021), suggesting that persistence in the E.A. has completely vanished during the last 7 years or so. Thus, when $\gamma$ is used to measure persistence both tests suggest strong evidence of a break in persistence following the launching of the euro as the common currency, and the implementation of single monetary policy by the European Central Bank for the twelve euro area countries.

Of course, whether the strong reduction in persistence in the E.A., as detected by $\gamma$, is a direct consequence of the implementation of a common monetary policy by the ECB through the anchoring of inflation expectations, remains an open issue. In fact, the possibility of such a decline stemming from a change in the type of shocks that hit
the E.A. during the last seven years or so, may not be ruled out on a priori grounds. But such an investigation is outside the scope of the present paper.

6 Conclusions

This paper suggests a new scalar measure of persistence, denoted by $\gamma$, together with a companion estimator, $\hat{\gamma}$, which explores the relationship between persistence and mean reversion. The new measure is defined as the unconditional probability of a stationary stochastic process not crossing its mean in period $t$, and has the property of being model free as its use does not require the specification and estimation of a model. Moreover, it is broader in scope than other measures used in the literature, in particular, than the widely used $\rho$, the sum of the coefficients of a pure autoregressive process.

It is shown that $\hat{\gamma}$ is an unbiased estimator of $\gamma$ when the mean of the process is known and a consistent estimator of $\gamma$ when the mean is unknown. Moreover, $\hat{\gamma}$ is asymptotically normal distributed. Inference on $\gamma$ may be conducted resorting either to the conventional approach in which a consistent kernel estimator for the asymptotic variance of $\hat{\gamma}$, $\sigma^2_{\hat{\gamma}}$, is used, or to the approach suggested in Kiefer and Vogelsang (2002), in which a non-consistent kernel estimator for $\sigma^2_{\hat{\gamma}}$ is used.

Using Monte Carlo simulations the finite sample properties of $\hat{\gamma}$ are compared to those of $\hat{\rho}$, the OLS estimator of $\rho$. We find that $\hat{\gamma}$, which by construction is immune to potential model misspecifications, is also robust to the presence of outliers in the data and that the coverage ratio of their empirical confidence intervals compares favourably to that of $\hat{\rho}$, when inference on $\hat{\gamma}$ is conducted using the approach suggested in Kiefer and Vogelsang (2002). However, in empirical applications the use of this approach, as a valid alternative to conventional methods, must be weighted against the implied loss in the power of the tests.

The relation between $\gamma$ and some alternative measures of persistence available in the literature, namely $\rho$, is also analysed. It is shown that there is a monotonic relationship
between $\rho$ and $\gamma$ when the data is generated by an AR(1) process, but for higher order processes the different scalar measures of persistence may deliver conflicting views on the degree of persistence. For that reason the use of $\gamma$ and $\rho$ as companion measures of persistence is suggested, as a strategy that may allow uncovering important features of the underlying process, that would otherwise remain unidentified.

Finally, the use of the new measure of persistence is illustrated by evaluating inflation persistence in the United States and the Euro Area for the period 1984-2008. In line with other empirical studies for a similar period, we find that, conditional on a break in the mean of inflation, both countries exhibit low levels of persistence and there is no evidence that persistence has changed with the average level of inflation. However, when we look for changes in persistence not related to changes in the average level of inflation, the use of $\gamma$ allows us to detect a significant reduction in persistence in the Euro Area, occurring around 2001/2002, soon after the launch of the euro and the implementation of a common monetary policy by the European Central Bank. This reduction, which would have remained undetected if we stuck to $\rho$ as the single measure of persistence, is such that persistence in the Euro Area becomes virtually nil for the period 2002-2008.
References


Table 1
Comparing $\gamma$ with other measures of persistence
AR(1) case

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\gamma$</th>
<th>$h$</th>
<th>$m_{50}$</th>
<th>$m_{95}$</th>
<th>$m_{99}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>0.00</td>
<td>0.500</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0.20</td>
<td>0.564</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0.40</td>
<td>0.631</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>0.60</td>
<td>0.705</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>0.70</td>
<td>0.747</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>0.80</td>
<td>0.795</td>
<td>3</td>
<td>3</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>0.90</td>
<td>0.856</td>
<td>6</td>
<td>6</td>
<td>28</td>
<td>43</td>
</tr>
<tr>
<td>0.95</td>
<td>0.899</td>
<td>13</td>
<td>13</td>
<td>58</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 2
Comparing $\gamma$ with other measures of persistence
Different AR(2) models with $\rho = \rho_1 + \rho_2 = 0.80$

<table>
<thead>
<tr>
<th>Model</th>
<th>$\gamma$</th>
<th>$h$</th>
<th>lar</th>
<th>$m_{50}$</th>
<th>$m_{95}$</th>
<th>$m_{99}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>$\rho_1 = 1.7, \rho_2 = -0.9$</td>
<td>0.853</td>
<td>5</td>
<td>–</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(2)</td>
<td>$\rho_1 = 1.5, \rho_2 = -0.7$</td>
<td>0.844</td>
<td>4</td>
<td>–</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(3)</td>
<td>$\rho_1 = 1.2, \rho_2 = -0.4$</td>
<td>0.828</td>
<td>4</td>
<td>–</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(4)</td>
<td>$\rho_1 = 1.0, \rho_2 = -0.2$</td>
<td>0.814</td>
<td>3</td>
<td>0.724</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>(5)</td>
<td>$\rho_1 = 0.8, \rho_2 = 0.0$</td>
<td>0.795</td>
<td>3</td>
<td>0.800</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>(6)</td>
<td>$\rho_1 = 0.6, \rho_2 = 0.2$</td>
<td>0.770</td>
<td>2</td>
<td>0.839</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>(7)</td>
<td>$\rho_1 = 0.4, \rho_2 = 0.4$</td>
<td>0.732</td>
<td>1</td>
<td>0.863</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>(8)</td>
<td>$\rho_1 = 0.2, \rho_2 = 0.6$</td>
<td>0.667</td>
<td>1</td>
<td>0.881</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>(9)</td>
<td>$\rho_1 = 0.0, \rho_2 = 0.8$</td>
<td>0.500</td>
<td>1</td>
<td>0.894</td>
<td>6</td>
<td>26</td>
</tr>
</tbody>
</table>
Table 3
Monte Carlo simulations – AR(1) model (T=100)

<table>
<thead>
<tr>
<th>True ρ</th>
<th>True γ</th>
<th>Model with no intercept</th>
<th>Model with an intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>True Mean bias of ρ (%)</td>
<td>True Mean bias of γ (%)</td>
</tr>
<tr>
<td>0.00</td>
<td>0.500</td>
<td>0.500 -0.010</td>
<td>0.497 -0.60</td>
</tr>
<tr>
<td>0.20</td>
<td>0.564</td>
<td>0.564 -7.78</td>
<td>0.560 -0.71</td>
</tr>
<tr>
<td>0.40</td>
<td>0.631</td>
<td>0.631 -5.41</td>
<td>0.626 -0.80</td>
</tr>
<tr>
<td>0.60</td>
<td>0.705</td>
<td>0.705 -4.66</td>
<td>0.698 -1.00</td>
</tr>
<tr>
<td>0.70</td>
<td>0.747</td>
<td>0.747 -4.47</td>
<td>0.739 -1.07</td>
</tr>
<tr>
<td>0.80</td>
<td>0.795</td>
<td>0.796 -4.36</td>
<td>0.786 -1.13</td>
</tr>
<tr>
<td>0.90</td>
<td>0.856</td>
<td>0.857 -4.37</td>
<td>0.842 -1.64</td>
</tr>
<tr>
<td>0.95</td>
<td>0.899</td>
<td>0.899 -4.52</td>
<td>0.878 -2.34</td>
</tr>
</tbody>
</table>

Table 4
Different AR(2) models with ρ = ρ₁ + ρ₂ = 0.80 (T=100)

<table>
<thead>
<tr>
<th>Model</th>
<th>Model with no intercept</th>
<th>Model with an intercept</th>
<th>True Mean</th>
<th>Estimated Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.800 0.00 0.797 -0.38 0.853 0.853 0.852 -0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.796 -0.50 0.789 -1.38 0.844 0.844 0.842 -0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.790 -1.25 0.777 -2.88 0.828 0.828 0.823 -0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.786 -1.75 0.769 -3.88 0.814 0.814 0.806 -0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.783 -2.13 0.760 -5.00 0.795 0.796 0.786 -1.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.779 -2.63 0.751 -6.13 0.770 0.771 0.758 -1.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.775 -3.13 0.743 -7.13 0.732 0.733 0.716 -2.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.772 -3.50 0.734 -8.25 0.667 0.667 0.645 -3.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.768 -4.00 0.725 -9.38 0.500 0.499 0.471 -5.80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

39
### Table 5
Robustness to additive outliers

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\gamma$</th>
<th>$\bar{\rho}$</th>
<th>Bias of $\hat{\rho}$ (%)</th>
<th>Inter Quartile Range of $\hat{\rho}$</th>
<th>$\bar{\gamma}$</th>
<th>Bias of $\hat{\gamma}$ (%)</th>
<th>Inter Quartile Range of $\hat{\gamma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>0.00</td>
<td>0.500</td>
<td>0.002</td>
<td>—</td>
<td>-0.059-0.063</td>
<td>0.500</td>
<td>0.14</td>
<td>0.47-0.53</td>
</tr>
<tr>
<td>0.20</td>
<td>0.564</td>
<td>0.105</td>
<td>-46.59</td>
<td>0.036-0.172</td>
<td>0.559</td>
<td>-0.87</td>
<td>0.53-0.59</td>
</tr>
<tr>
<td>0.40</td>
<td>0.631</td>
<td>0.219</td>
<td>-44.34</td>
<td>0.134-0.302</td>
<td>0.621</td>
<td>-1.60</td>
<td>0.59-0.66</td>
</tr>
<tr>
<td>0.60</td>
<td>0.705</td>
<td>0.359</td>
<td>-39.07</td>
<td>0.257-0.463</td>
<td>0.689</td>
<td>-2.30</td>
<td>0.65-0.72</td>
</tr>
<tr>
<td>0.70</td>
<td>0.747</td>
<td>0.448</td>
<td>-34.76</td>
<td>0.342-0.561</td>
<td>0.729</td>
<td>-2.50</td>
<td>0.69-0.76</td>
</tr>
<tr>
<td>0.80</td>
<td>0.795</td>
<td>0.560</td>
<td>-28.62</td>
<td>0.459-0.677</td>
<td>0.774</td>
<td>-2.77</td>
<td>0.74-0.81</td>
</tr>
<tr>
<td>0.90</td>
<td>0.856</td>
<td>0.713</td>
<td>-19.33</td>
<td>0.628-0.818</td>
<td>0.824</td>
<td>-2.71</td>
<td>0.80-0.87</td>
</tr>
<tr>
<td>0.95</td>
<td>0.899</td>
<td>0.815</td>
<td>-12.61</td>
<td>0.755-0.903</td>
<td>0.877</td>
<td>-2.46</td>
<td>0.84-0.91</td>
</tr>
</tbody>
</table>

### Table 6
Coverage ratio of 95% confidence intervals for $\rho$ and $\gamma$

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\gamma$</th>
<th>Coverage ratio for $\rho$</th>
<th>Coverage ratio for $\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T=100</td>
<td>T=250</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>0.00</td>
<td>0.500</td>
<td>0.951</td>
<td>0.951</td>
</tr>
<tr>
<td>0.20</td>
<td>0.564</td>
<td>0.949</td>
<td>0.951</td>
</tr>
<tr>
<td>0.40</td>
<td>0.631</td>
<td>0.946</td>
<td>0.950</td>
</tr>
<tr>
<td>0.60</td>
<td>0.705</td>
<td>0.942</td>
<td>0.948</td>
</tr>
<tr>
<td>0.70</td>
<td>0.747</td>
<td>0.938</td>
<td>0.946</td>
</tr>
<tr>
<td>0.80</td>
<td>0.795</td>
<td>0.935</td>
<td>0.943</td>
</tr>
<tr>
<td>0.90</td>
<td>0.856</td>
<td>0.922</td>
<td>0.937</td>
</tr>
<tr>
<td>0.95</td>
<td>0.899</td>
<td>0.885</td>
<td>0.921</td>
</tr>
</tbody>
</table>
### Table 7
Coverage ratio of 95% confidence intervals for $\rho$ and $\gamma$

AR(2) models with an intercept and $\rho_1 + \rho_2 = 0.80$

<table>
<thead>
<tr>
<th>AR(2) Models</th>
<th>True Coverage ratio for $\rho$</th>
<th>Coverage ratio for $\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K=8 Andrews</td>
<td>KV</td>
</tr>
<tr>
<td></td>
<td>T=100 T=250</td>
<td>T=100 T=250</td>
</tr>
<tr>
<td>(2) (3) (4)</td>
<td>(5) (6) (7) (8) (9) (10)</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.853 0.947 0.952</td>
<td>0.983 0.985</td>
</tr>
<tr>
<td>(2)</td>
<td>0.844 0.945 0.950</td>
<td>0.946 0.953</td>
</tr>
<tr>
<td>(3)</td>
<td>0.828 0.937 0.946</td>
<td>0.918 0.936</td>
</tr>
<tr>
<td>(4)</td>
<td>0.814 0.933 0.944</td>
<td>0.910 0.930</td>
</tr>
<tr>
<td>(5)</td>
<td>0.795 0.928 0.943</td>
<td>0.895 0.923</td>
</tr>
<tr>
<td>(6)</td>
<td>0.770 0.923 0.943</td>
<td>0.887 0.917</td>
</tr>
<tr>
<td>(7)</td>
<td>0.732 0.921 0.940</td>
<td>0.875 0.907</td>
</tr>
<tr>
<td>(8)</td>
<td>0.667 0.917 0.939</td>
<td>0.856 0.896</td>
</tr>
<tr>
<td>(9)</td>
<td>0.500 0.913 0.938</td>
<td>0.815 0.867</td>
</tr>
</tbody>
</table>

*(a) The AR(2) models are the same as in Table 2;*

### Table 8
Inflation persistence in the U.S. and the Euro Area

<table>
<thead>
<tr>
<th></th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No break in the mean</td>
</tr>
<tr>
<td></td>
<td>full sample estimates(*)</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>$\hat{\gamma} = 0.646$</td>
<td>$\hat{\gamma} = 0.545$</td>
</tr>
<tr>
<td>(0.048)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>$\hat{\rho} = 0.749$</td>
<td>$\hat{\rho} = 0.666$</td>
</tr>
<tr>
<td>(0.113)</td>
<td>(0.142)</td>
</tr>
<tr>
<td></td>
<td>Euro Area</td>
</tr>
<tr>
<td>$\hat{\gamma} = 0.788$</td>
<td>$\hat{\gamma} = 0.677$</td>
</tr>
<tr>
<td>(0.026)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>$\hat{\rho} = 0.866$</td>
<td>$\hat{\rho} = 0.526$</td>
</tr>
<tr>
<td>(0.065)</td>
<td>(0.128)</td>
</tr>
</tbody>
</table>

(*) For $\hat{\gamma}$ and $\hat{\alpha}_2$ KV (2002) standard-errors are given in parentheses; $\hat{\alpha}_2 = \hat{\gamma}_1 - \hat{\gamma}_2$ and $\hat{\lambda} = \hat{\rho}_2 - \hat{\rho}_1$ where $\hat{\gamma}_1$ and $\hat{\rho}_1$ refer to the first, and $\hat{\gamma}_2$ and $\hat{\rho}_2$ to the second sub-period, respectively.

41
Graph No.1 - Realization of models 1, 5 and 9.

Graph No.2 - Quarter-on-quarter U.S. and E.A. inflation (no break in the mean).
Graph No.3 - Quarter-on-quarter U.S. inflation (break in the mean).

Graph No.4 - Quarter-on-quarter E.A. inflation (break in the mean).