Convergence of Euro Area Inflation Rates

Claude Lopez [*]
Banque de France

David H. Papell [**]
University of Houston

December 14, 2011

Abstract

We study the behavior of inflation rates among the 12 initial Euro countries in order to test whether and when the group convergence initially dictated by the Maastricht treaty and now by the ECB, occurs. We also assess the impact of events such as the advent of the Euro and the 2008 financial crisis. Due to the small size of the estimation sample, we propose a new procedure that increases the power of panel unit root tests when used to study group-wise convergence. Applying this new procedure to Euro area inflation, we find strong and lasting evidence of convergence among the inflation rates soon after the implementation of the Maastricht treaty and a dramatic decrease in the persistence of the differential after the occurrence of the single currency. After the 2008 crisis, Euro area inflation rates follow the ECB's price stability benchmark, although Greece reports relatively higher inflation.

JEL: C32, E31

Keywords: groupwise convergence, inflation, Euro area

1. Introduction

Inflation rates and their convergence within the Euro area have been a major concern since well before the advent of the single currency. The recent financial crisis and its strong impact on several Euro area countries with higher inflation rates have strengthened this interest, especially in the light of the European Central Bank's (ECB) objective of price stability. Since 1999, the ECB has clearly stated that Euro Area inflation should be maintained below but close to 2% in the medium run.

Assessing whether Euro area inflation rates satisfy such a convergence since the advent of the single currency is quite challenging due to the limited amount of available data and the poor performance of standard time series techniques typically used to test whether convergence has occurred. We address this issue in two steps. First, we investigate group-wise convergence among Euro area inflation rates using an improved testing procedure and focusing on the 12 initial members which share a common Euro period. Then, we look at whether these inflation rates respect the ECB's benchmark, both as a group and individually.

Time series investigation of the convergence hypothesis often relies on unit root tests. The rejection of the null hypothesis is commonly interpreted as evidence that the series *have converged* to their equilibrium state, since any shock that causes deviations from equilibrium eventually dies out. The extension of these tests to the panel framework has significantly influenced the literature on how to measure convergence of macroeconomic variables.

Panel unit root tests for convergence among series, or group-wise convergence, utilize Bernard and Durlauf's (1995, 1996) definition of time series convergence for long-run output movements, where two (or more) countries *have converged* when long-run forecasts of per capita output differences tend to zero as the forecasting horizon tends to infinity. In the bivariate context, tests for time series convergence require cross-country per capita output differences to be stationary.[1] In the multivariate or panel context, a group of countries have converged if the null hypothesis that the difference between each country's output and the cross-sectional mean has a unit root can be rejected in favor of the alternative hypothesis that each difference is stationary.[2] Several works use panel methods to investigate output convergence (Ben-David (1993, 1996), Islam (1995), Evans and Karras (1996), Evans (1998),

---

[1] Clearly, the time series' definition of convergence implies that the series have reached the state of equilibrium and not that they are adjusting toward it.

[2] Pesaran(2007) suggests an alternative to test for pairwise convergence of ouput.

1

and Fleissig and Strauss (2001), among others) or inflation convergence (Lee and Wu (2001), Kočenda and Papell (1997) and Weber and Beck (2005), among others).

While panel unit root tests have significantly enhanced finite sample performance when compared to the univariate approach, these tests can still have low power to reject the unit root null in a panel of stationary series if the panels consist of highly persistent series, contain a small number of series, and/or have series with a limited length. This paper proposes a new procedure that improves the finite sample power of panel unit root tests when testing for group-wise convergence and uses the procedure to analyze the behavior of the inflation rates across the Euro area, isolating prior and Euro periods.

The suggested method uses information known prior to estimation. Panel unit root tests for group-wise convergence involve stationarity between a group of series and their cross-sectional means. As the series may not be characterized by absolute convergence toward the cross-sectional average, each differential can have a non-zero mean. By construction, however, the group of differentials has a cross-sectional average of zero for each time period. In order to improve the panel unit root test's performance, we exploit this extra information on the data by incorporating the appropriate restriction when estimating the model and generating finite sample critical values. Monte Carlo simulations confirm the enhanced finite sample performance of the test when using the constraint. To our knowledge, this constraint has not been utilized for previous tests of convergence using panel unit root tests.[3]

The improved performance in small samples allows us to focus on a smaller data span while maintaining the reliability of the analysis. In particular, it enables us to isolate the estimation period from 1999:1 to 2006:12, starting with the advent of the Euro and ending before the 2008 financial crisis. The study analyzes Euro area inflation rates from 1979:1 to 2010:04 by constructing a rolling window of eight years, starting with 1979:1-1986:12 and ending with 2002:5-2010:4. The windows starting in 1999:1 account for the Euro period while the windows starting in 2001:1 include the 2008 financial crisis. This rolling window approach also deals with any potential time break in the data due to events such as German reunification.

---

[3] It should be emphasized that our proposed method is only applicable for tests of group-wise convergence. The power of panel unit root tests that examine the Purchasing Power Parity hypothesis by investigating the stationarity of real exchange rates, for example, cannot be improved by our method as, in this case, the series are individually converging to their own mean but not to a common target.

The first part of our analysis shows that inflation rates within the Euro area share a common pattern: they have converged or report a stable relation among themselves as early as just after the implementation of the Maastricht treaty and this convergence remains strong until after the advent of the Euro. The results also highlight the temporary impact of the 2008 financial crisis: an initial weakening of the unit root rejections is shortly followed by strengthened evidence of convergence for most of the groups observed. The rate of persistence of the differentials, which is directly linked to the degree of convergence among the inflation rates, highlights four phases: (i) periods ending between 1986:12 and 1997:8, where the Maastricht criteria are not fully implemented and the persistence is quite high but stable; (ii) periods ending between 1997:9 and 2004:12, which is a period of transition from implementing the Maastricht treaty to the advent of the Euro, where the persistence varies a lot, with an initial drastic decrease that is later partially compensated; (iii) periods ending between 2005:12 and 2007:12, the Euro period, where the persistence is, once again, stable, yet at a lower level than in the initial period and, finally, periods ending after 2008:09 that include the 2008 financial crisis.[4] The described behavior follows closely the European Monetary Union time table. The generated median unbiased estimates, their 95% confidence intervals and the corresponding half-lives confirm a dramatic decrease in the persistence of the differentials after the occurrence of the single currency.

The second part of the analysis focuses on the ECB's objective and on individual deviations from it. The results confirm that the unique currency significantly anchors the individual and cross-sectional volatility of the Euro area inflation rates, especially for less wealthy countries such as Greece, Portugal, and Spain.[5] Euro area inflation is very close to 2% for most of the Euro period with the exception of the 2008 financial crisis. However, at the country level, Ireland, Greece, and Spain relax their efforts shortly after the adoption of the single currency, ending up with noticeably higher inflation rates than the other Euro countries, placing them in a strong disadvantage in term of price competitiveness and susceptible to current account deficits and bubbles (real estate: Spain, Ireland, public debt: Greece). While these three countries are the most affected among the group by the crisis, this

---

[4] While the euro started in 1999, fixed parities among the currencies were set in 1998.
[5] The same question cannot be investigated for Ireland as its monthly data starts only in 1999.

event seems to act as a "realignment" for their inflation with the rest of the group for Spain and Ireland, but not for Greece.

2. Panel Unit Root Tests for Convergence

2.1 Group-wise stochastic convergence

In the panel framework, testing for (stochastic) convergence of a group of N time series requires studying the dynamic properties of the series differential with respect to the cross sectional mean. Group-wise (stochastic) convergence implies that[6]:

$$\lim_{k \to \infty} E\left(y_{i,t+k} - \sum_{j=1}^{N} \frac{y_{j,t+k}}{N} \,\Big|\, I_t \right) = \eta_i \; for \; i = 1, ..., N \tag{1}$$

Where I_t represents the information set available at time t. If $\eta_i = 0$, the convergence follows Bernard and Durlauf (1996)'s definition of absolute convergence. If $\eta_i \neq 0$ the convergence is said to be conditional or relative as defined by Durlauf and Quah (1999), which implies that the series *have converged* toward a time-invariant equilibrium differential. [7]

2.2 Panel unit root test

We modify standard panel unit root tests to account for the restriction on the intercepts when testing for group-wise convergence. More specifically, we focus on the second generation of panel unit root tests that account for contemporaneous correlation by estimating the residual covariance matrix. The test considered is an extension of the Levin, Lin, and Chu (2002) application of the ADF test to the panel framework that investigates a homogeneous rate of convergence across the series. Let consider the following system of ADF regressions:

$$\Delta y_{it} = c_i + \rho_i \, y_{i,t-1} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j} + \varepsilon_{it} \tag{2}$$

$$with \quad i = 1,...,N \quad t = 1,...,T \quad and \quad \varepsilon_{it} \sim WN(0;\Sigma)$$

---

[6] The notion of group-wise convergence refers to convergence toward a common behavior captured by the cross-sectional mean. Looking at the convergence of the inflation rate of each country toward a cross-sectional mean that excludes the country in question does not address this issue.

[7] The differentials will be stationary if either the series and the cross-sectional mean are both I(0) or if they are both I(1) and cointegrated.

where $\rho_i = \rho$ is the homogeneous rate of convergence, k_j the lagged first differences that account for serial correlation and $\sum$ the non-diagonal covariance matrix. The null and alternative hypotheses tested are $\rho = 0$ and $\rho < 0$.

The pooled ADF test relies on feasible generalized least squares (SUR) method, hence the name ADF-SUR test. It is performed in two steps. First, for each series, k_j is selected with the recursive lag-selection procedure of Hall (1994). Then, the residuals covariance matrix is deduced and used to estimate (2) with the SUR method, constraining the values of ρ to be identical across equations and using k_j. Finally, the estimated ρ and its corresponding standard deviation allow us to calculate the t- statistics corresponding to the null $\rho = 0$. Since the focus of the paper is on a panel of macroeconomic variables where the time series dimension is large compared to the cross-section dimension, it is assumed that $T > N$.

While it would be desirable to allow for heterogeneous rates of convergence, the choices are problematic.[8] Following Im, Pesaran, and Shin (2003), several tests that average t-statistics across the members of the panel have been developed.[9] The alternative hypothesis for these tests, however, is that $\rho_i < 0$ for at least one i, which is not economically relevant for investigating convergence among a group of countries. The tests developed by Breuer, McNown, and Wallace (2002), which allow ρ_i to be heterogeneous across countries in a framework similar to (2), provides (at best) modest increases in power over univariate tests.

2.3 The new testing procedure

Our testing procedure benefits from extra knowledge available about the data and designs a model that accounts for all information available prior to the estimation. More specifically, this non-sample information is included as a restriction in the estimation and when generating the finite sample critical values. The restriction being true by construction, the final estimator ends up with a smaller variance than the unrestricted one. Greene (2008,

---

[8] Breitung and Pesaran (2005) survey the existing literature and point out that, in both the homogenous and the heterogeneous cases, the rejection of the null hypothesis means that "a significant fraction of the cross-section units is stationary".

[9] An alternative is to use a factor structure approach as in Bai and Ng (2004).

p89) suggests that "one way to interpret this reduction in variance is as the value of the information contained in the restriction". [10]

The procedure relies on the knowledge that, once transformed, the data may have a non-zero mean for each differential i but a cross-sectional mean equal to 0 at every period. If y_{it}^{diff} is the differential for country i at time t with respect to the cross-sectional mean such that $y_{it}^{diff} = y_{it} - \sum_{i=1}^{N} y_{it} / N$; then by construction, for each period of time t, the sum of the differentials is equal to 0, that is $\sum_{i=1}^{N} y_{it}^{diff} = \sum_{i=1}^{N} \left(y_{it} - \frac{\sum_{i=1}^{N} y_{it}}{N} \right) = 0$. Let replace y_{it} by y_{it}^{diff} in (2), then the intercepts c_i are on average equal to 0. Hence, the estimation uses the restriction $\sum_{i=1}^{N} c_i = 0$. Note that, since each regression allows for an intercept, we are not testing for absolute convergence. The resulting system of equations is:

$$\Delta y_{it}^{diff} = c_i + \rho y_{it-1}^{diff} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{it-j}^{diff} + \varepsilon_{it}$$

$$\text{with} \quad i = 1,...,N \quad t = 1,...,T \quad and \quad \varepsilon_{it} \sim WN$$

(3)

where ρ is the homogenous rate convergence, and y_{it}^{diff} is the data differential with respect to the cross-sectional mean. The error terms $(\varepsilon_{1t},...,\varepsilon_{Nt})$ are stationary with a non-diagonal covariance matrix $\sum$. The standard hypotheses, H$_0$: $\rho = 0$ versus H$_1$: $\rho < 0$, are tested.

The estimation procedure follows three steps:

1. *Data transformation*: the differentials with respect to the cross-sectional mean are calculated for all series

2. *Lag selection*: the number of lagged first difference terms allowing for serial correlation, k_i in (3), is selected using the recursive procedure for each series

3. *Estimation*: The residual covariance matrix $\sum$ is estimated. The resulting $\hat{\Sigma}$, along with the pre-selected k_i, is then used in the estimation of (3) with the SUR method while two restrictions are imposed:

---

[10] Judge et al. (1988, p812) explains that" if nonsample information is correct, then using it in conjunction with the sample information will lead to an unbiased estimator that has a precision matrix superior to the unrestricted least squares estimator".

a. $\sum_{i=1}^{N} c_i = 0$, that is the non-sample information

b. $\rho_i = \rho$, that is a homogeneous rate of convergence

The estimated ρ and its corresponding standard deviation are obtained, and the t-statistic is calculated for $H_0 : \rho = 0$.

The interpretation of the two restrictions is very different. (a) is true by construction, and therefore there is no question whether or not it is correct. (b) is almost surely false, as there is no reason why each country should have the same rate of convergence. There are two ways, however, for the restriction of homogeneous convergence rates to be false. First, all of the $\rho_i < 0$. In that case, rejection of the unit root null correctly provides evidence of convergence. Second, some of the $\rho_i < 0$ and some of the $\rho_i = 0$. In that case, there is a mixed panel and rejection of the unit root null does not correctly provide evidence of convergence. We consider the performance of our test with mixed panels below.

O'Connell (1998), Maddala and Wu (1999), and Lopez (2009a), among others, show that panel unit root tests estimating the residual covariance matrix should rely on simulated critical values to reduce any size distortions due to the cross-sectional correlation, while Chang (2004) proves the asymptotic validity of a sieve bootstrap procedure for non-pivotal homogeneous panel unit root tests. As a result, the bootstrap critical values are generated using the following non-parametric resampling method with replacement: first, the bootstrap innovations u_{it}^{*} are obtained by resampling with replacement the empirical residuals estimated

from $\Delta y_{it}^{diff} = \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j}^{diff} + u_{it}$.[11] The contemporaneous correlation is preserved by resampling

the residuals $\left(u_t^{*} \right)$ as a vector. Next, the bootstrap samples $\left(\varepsilon_{it}^{*} \right)$ are recursively generated using

the estimated parameters $(\hat{\phi}_{ij})$ and the bootstrap samples $\left(u_t^{*} \right)$ as $\varepsilon_{it}^{*} = \sum_{j=1}^{k_i} \hat{\phi}_{ij} \varepsilon_{i,t-i}^{*} + u_{it}^{*}$, starting

from $\left(u_{i0}, \dots, u_{i,-k_i+1} \right)$. Finally, the pseudo-data $y_{it}^{diff\,*}$ are obtained by taking the partial sum of

---

[11] The empirical residuals were, first, centered then resampled.

$\left(\varepsilon_{it}^{*}\right)$ as $y_{it}^{diff*} = y_{i0}^{diff*} + \sum_{j=1}^{k_i} \varepsilon_{ij}^{*}$.[12] The estimation procedure explained in Section 2.3 is then applied and the t-statistic calculated. This procedure is iterated 1000 times, the resulting vector of t-statistics is then sorted to calculate either the data specific critical values or the p-values. Each estimation requires its own set of critical values to be generated.

Davidson and G. MacKinnon (2006) explain that "imposing the restriction […] yields more efficient estimates of the nuisance parameters upon which the distribution of the test statistics may depend. This generally makes bootstrap test more reliable, because the parameters of the bootstrap DGP are estimated more precisely". Since the restriction is true by construction, we expect the restricted test to perform better in small samples than the unrestricted one.

2.4 Impact of the constraint in small samples

In order to analyze the impact of the restriction $\sum_{i=1}^{N} c_i = 0$, a set of simulations investigates the finite sample performance of the ADF-SUR test with and without the restriction. Let consider the following data generating processes:

$$y_{it} = \rho y_{i,t-1} + u_{i,t} \qquad \text{with } i=1, \dots, N \text{ and } t=1, \dots, T$$

The innovations $\{u_{it}\}$ are drawn from *iid* normal distributions with mean zero and a diagonal covariance matrix Σ.[13] The panel dimensions are $N = 5$, 10, and 20 and $T = 25$, 50, 100, and 200. For each experiment, the finite sample critical values and the empirical rejection probabilities calculated at a 5% nominal level are based on 2000 iterations.[14,15] Since we are using randomly generated data, each experiment is repeated 20 times, hence Tables 1 and 2 report the average rejection probabilities.

Table 1 reports the finite sample properties of both restricted and unrestricted ADF-SUR tests. The data sets are generated under the null hypothesis ($\rho = 1.00$) for the size and

---

[12] Each pseudo-data $\left(y_{it}^{diff*}\right)$ is generated with T+50 observations, then the first 50 observations are discarded, hence each $\left(y_{i0}^{diff*}\right)$ is random.

[13] Similar simulations have been reproduced using non-diagonal matrix covariance, that is including and accounting for contemporaneous correlation, without any significant change regarding the impact of the restriction on the intercept.

[14] Davidson and McKinnon (1999) advise a minimum of 1500 bootstraps when analyzing the performance of the test at 1%.

[15] Davidson and McKinnon (2006) define and discuss this probability for the power and size of bootstrap tests

under the alternative (ρ = 0.99, 0.97, 0.95 and 0.90) for the size adjusted power.[16] Both tests report almost no size distortion with a probability of rejecting the unit root null when the data have one, close the nominal size of 5%. However, the tests significantly differ in their ability of rejecting accurately the null hypothesis when analyzing stationary data. For example, for highly persistent data such that (N, T, ρ) = (10, 100, 0.97), the restriction increases the size-adjusted power of the ADF-SUR test from 0.384 to 0.595. Similarly, for moderately persistent data such that (N, T, ρ) = (20, 50, 0.95), the restriction increases the power from 0.337 to 0.539. As expected, these improvements disappear as N and T increase and the data is less persistent, that is in the cases where the ADF-SUR test performs well. In addition, the restriction has only a moderate impact when the panel has a small time dimension, $T = 25$ and 50, and the data is extremely persistent, $\rho = 0.99$. In sum, the restriction significantly enhances the test's performance for persistent data ($\rho > 0.9$) and small to medium data spans ($T < 200$).

Table 2 focuses on the test's performance when the data is not generated under the alternative hypothesis of homogeneous and stationary rates of convergence but as a mix of stationary and non-stationary processes. More specifically, some series converge at a same rate ($\rho_i = \rho = 0.97$, 0.95, 0.90 and 0.8 for $i = 1,...,k$) while others follow a non-stationary process ($\rho_j = 1.0$ for $j = k+1, ..., N$).[17] The data length T is equal to 100 for $N = 5$, 10 and 20.

Such an experiment allows us to investigate whether the improved finite sample performance of the restricted test leads to an increase in unwanted rejections of the null hypothesis over the unrestricted test. Indeed, Taylor and Sarno (1998) and Breuer, McNown, and Wallace (2001) have provided evidence that, in the general case where the sum of the intercepts is not constrained to equal zero, the unit root null can be rejected by panel methods with homogeneous rates of convergence even when the panels contain only a few stationary series. Breuer, McNown, and Wallace (2002), Sarno and Taylor (2002), and Taylor and Taylor (2004) go further, arguing that the unit root null can be rejected even if only one of the series is stationary. To address this concern, we first look at the bottom row of Table 2, for N = 5, 10, and 20, that reports the (correctly sized 0.05) rejection frequencies when all series have a unit root. Going up one row, the rejection frequencies for both the restricted and the

---

[16] The case ρ = 0.8 is not reported as it does not provide any new insights on the test's behavior.
[17] The case ρ = 0.99 is not reported as it does not provide any new insights on the test's behavior.

unrestricted tests are depicted when one of the series is stationary, that is $(\rho_i, \rho_j) = (\rho, 1.00)$ for $i = 1$ and $j = 2,...,$N. For $N = 5$, they range from 0.07 ($\rho = 0.97$) to 0.11 ($\rho = 0.8$), for $N = 10$, they range from 0.06 ($\rho = 0.97$) to 0.08 ($\rho = 0.8$) and, for $N = 20$, they range from 0.06 ($\rho = 0.97$) to 0.07 ($\rho = 0.8$). Hence, it seems very unlikely that the inclusion of one stationary series will produce a rejection of the unit root null with any of these tests.[18]

While the argument that inclusion of one stationary series will produce rejections using panel unit root tests with homogeneous rates of convergence seems overstated, the results confirm that one needs to be careful about interpreting rejections of the null as evidence that all of the series are stationary. For example, with $N = 10$, both tests report a rejection frequency of about 0.50 with 8 stationary series if $\rho = 0.95$. Since the result of rejection or non-rejection would be analogous to the outcome of a coin flip, one would not want to conclude in favor or against the null hypothesis.

Yet, it is worth noting that, for all three panels with a mix of unit root and less persistent ($\rho = 0.9$ and 0.8) stationary series, the rejection frequencies for the restricted test are smaller than those for the unrestricted test. Hence, one would be less likely to falsely reject the unit root null hypothesis for most of the cases when using the restricted ADF-SUR test. For the panels with a mix of unit root and more persistent ($\rho = 0.95$ and 0.97) stationary series, the rejection frequencies for the restricted tests are still smaller or equal to those for the unrestricted tests except in presence of very few (up to three depending the panel) unit roots.

In practice, however, one is much less likely to falsely reject the unit root null with restricted than with unrestricted ADF-SUR tests. This is because, with highly persistent processes and $N = 5$ or $N = 10$, the tests do not have much ability to reject the unit root null even when all of the series are stationary. Taking the most extreme example $(N, \rho) = (5, 0.97)$ for emphasis, the 5% size adjusted power is only 0.41 for the restricted test and 0.23 for the unrestricted test when all of the series are stationary. With one stationary series, the fact that the rejection frequency is larger for the restricted (0.22) than the unrestricted (0.16) test is unlikely to cause an inappropriate conclusion as the restricted test under rejects the null hypothesis around 80% of the time.

---

[18] Some of our rejection frequencies without the constraint are lower than in Breuer, McNown, and Wallace (2001) for identical panels. The differences appear to be due to their use of Levin, Lin, and Chu (2002) critical values which do not account for serial correlation. Papell (1997) discusses this issue.

A very different picture emerges with less persistent processes where the tests are often able to appropriately reject the unit root null when all of the series are stationary. We will focus on a comparison of the rejection frequencies between the two tests for the smallest number of stationary series for which the rejection frequency of the unrestricted test is 0.50 or higher. For $N = 5$, the rejection frequency is 0.58 for the restricted test and 0.65 for the unrestricted test with 4 stationary series and $\rho = 0.9$ and is 0.40 for the restricted test and 0.51 for the unrestricted test with 3 stationary series and $\rho = 0.8$. With $N = 10$, the rejection frequency is 0.57 for the restricted test and 0.66 for the unrestricted test with 7 stationary series and $\rho = 0.9$ and is 0.54 for the restricted test and 0.64 for the unrestricted test with 6 stationary series and $\rho = 0.8$. When $N = 20$, the rejection frequency is 0.46 for the restricted test and 0.56 for the unrestricted test with 11 stationary series and $\rho = 0.9$ and is 0.42 for the restricted test and 0.53 for the unrestricted test with 9 stationary series and $\rho = 0.8$. In the above examples, both tests very often reject the unit root null when all of the series are stationary, so they represent cases where it is plausible that the unit root null might be rejected with a mixture of stationary and non-stationary series. While other examples could be chosen, the pattern is clear. For mixed panels that contain less persistent stationary series with $\rho = 0.8$ or $\rho = 0.9$, one is less likely to mistakenly reject the unit root null hypothesis with the restricted than with the unrestricted tests.

When the data is, by construction, restricted so that the sum of the intercepts is equal to zero for each period, the gain in efficiency obtained by imposing the restriction in the estimation has two main impacts on the ADF-SUR test.[19] First, the more precise estimation and resulting bootstrap procedure leads to a more powerful size-adjusted test for the most commonly encountered panel dimensions in macroeconomics. Second, the rejection frequencies are generally smaller for mixed panels of stationary and non-stationary processes. Combining the results, the restriction improves the overall behavior of the test, enhancing its ability to correctly reject the unit root null hypothesis when all series are stationary and to correctly fail to reject the unit root null when a subset of the series are non-stationary.

3. Inflation convergence within the Euro Area

---

[19]The gain in efficiency refers to the more precise estimation that leads to smaller variance of the error terms.

In light of the achievement of the Maastricht criteria, the fixing of Euro area exchange rates in mid-1998, and the establishment of the Euro in January 1999, one would expect Euro area inflation rates to have converged during the period immediately preceding the advent of the single currency. This expectation is confirmed by numerous studies, including Rogers, Hufbauer and Wada (2001), Engel and Rogers (2004), Weber and Beck (2005), Busetti, Forni, Harvey and Venditti (2007) and Rogers (2007), which agree that prices were less dispersed and inflation rates among Euro area countries have converged in the mid-1990s. In contrast, research investigating the post-1998 period, including ECB (2003), Honohan and Lane (2003), Engel and Rogers (2004), Weber and Beck, (2005), Rogers (2007), and Fritsche and Kuzin (2008), concludes that the advent of the Euro resulted in the weakening of inflation convergence among the Euro area countries and in an increase in their price dispersion. An exception is Honohan and Lane (2004), who report sharp convergence in inflation rates since 2002.

The ECB's price-stability policy implies that "as a whole" the Euro area should have a yearly inflation equal to or less than 2%. If such a goal is achieved then Euro area inflation rates should have (group-wise) converged toward a cross-sectional mean of 2% or less.

Our study contributes towards understanding the behavior of the 12 initial Euro area inflations and their evolution through time by (i) investigating the group behavior of the inflation differentials with respect to their cross-sectional mean and (ii) looking at individual deviations from the group central tendency and from the ECB's benchmark. We first investigate the evolution of group-wise convergence over time, starting with the period prior to the European Monetary System and ending with the post 2008 financial crisis period. Next, we highlight the impact of the Euro by comparing the estimated speeds of convergence before and after the adoption of the single currency as well as the post 2008 financial crisis. Finally, we focus on the ECB's 2% benchmark and how each inflation rate compares to it.

3.1 Behavior of Euro area inflation rates as a group: groupwise convergence

Annual inflation rates with monthly data π_{it} for the ith country at time t are calculated such that: $\pi_{it} = \ln(CPI_t) - \ln(CPI_{t-12})$.[20] The differentials y_{it} are generated so that:

---

[20] While constructing yearly inflation rates from monthly data may create a moving average in the error term, McKinnon (2006) shows that bootstrapped tests perform well for data similar to ours. The data is seasonally adjusted.

$\pi_{it}^{diff} = \pi_{it} - \bar{\pi}_t$ where $\bar{\pi}_t$ is the cross-sectional average inflation rate.[21] Monthly CPI data are from International Financial Statistics from 1979:1 to 2010:4. Euro 10 (E10) countries are Austria, Belgium, Finland, France, Germany, Italy, Luxembourg, Netherlands, Portugal and Spain, Euro 11 (E11) countries include Greece while Euro 12 (E12) countries also include Ireland. [22]

Figure 1 reports the inflation rates as well as the cross-country means, medians and standard deviations of the inflation rates. The descriptive statistics show an overall decrease for both the E10 and E11 panels. More specifically, this decrease occurs in three phases: 1982-1987 has the steepest slope, followed by 1990-1999 with a flatter slope, and then 2000-early 2008 reports no visible change in the slope.[23] Finally, the last period, from mid-2008 to the end of the sample, presents some divergences between the behavior of the mean, the median and the standard deviation. For each group, the mean and the median show a noticeable increase, with the highest point at mid-2008; then a significant drop, with the lowest level in mid-2009 and concludes with a level lower than prior to the crisis. In contrast, the standard deviation for E10 and E11 remains quite stable since 2000, while E12 shows an increase in inflation dispersion since mid-2008. Note that the mean and the median for E10 and E11 are very close since 1997 and 2000, respectively. Only E12 reports a mean lower than the median for the year 2009. This increase in dispersion is coherent with the noticeable decrease in inflation of Ireland reported in Figure 4b. Finally, the E10, E11 and E12's cross sectional means are very close or below to the 2% benchmark for the periods 1997-2000, 2003-2007 and after 2009.

The enhanced performance of the new estimation procedure enables us to consider relatively short periods while preserving good size adjusted power of the test. We choose an estimation window of eight years as it corresponds to the Euro and prior-financial crisis period of 1999:1-2006:12 (96 monthly observations). The window is then rolled from 1979:1-1986:12 to 2002:5-2010:4, one month at a time, providing a detailed view of the adjustments that took place throughout the EMS changes. This approach also limits the impact of potential changes in the parameters on the estimation results while depicting the evolution of the results

---

[21] Yearly inflation with monthly data and annualized monthly average inflations yield to similar results.
[22] The monthly data for Ireland is available only starting 1998:1; hence the analysis based on E12 starts at that point.
[23] Lopez (2009b) shows that the Euro-zone inflation rates are regime-wise stationary.

through time. In contrast to studies which use a recursive (expanding) estimation window to study convergence, our results are not affected by the fact that the power of panel unit root tests increases with the number of observations as well as the size of the panel.

Figure 2 reports p-values for all panels using restricted and unrestricted ADF-SUR tests.[24] The three groups of countries lead to similar conclusions. A comparison between the restricted and the unrestricted estimations emphasizes the impact of the previously discussed gain in precision with the new estimation procedure: while the results observe a similar pattern, the restricted approach consistently leads to lower p-values.[25] The findings based on restricted estimation show sporadic rejections of the unit root hypothesis for pre-Euro windows ending in 1990-1994 and 1997-1998. These (lack of) results are important for two reasons. First, they emphasize how period specific the rejection of the unit root can be and the necessity of reporting lasting evidence of convergence when concluding in favor of stability. Second, these scarce rejections of the unit root are coherent with the troubles that EMS had during the 1980s.

The windows ending in 2000 – 2010 report the strongest evidence that convergence has occurred. Focusing on the restricted estimation, the unit root null is rejected (at least at the 10 percent level) with E10 for most of the windows ending in 2000:9 - 2010:4. [26]Adding Greece (E11) leads to relatively similar results, with rejection at least at 10 percent of the unit root for all the windows ending in 2002:3 - 2009:2 and after 2009:10. In contrast, the addition of Ireland (E12) has a noticeable impact as the unit root is almost never rejected after the window ending in 2009:2. The impact of imposing the restrictions is very clear for the Euro period, leading to lasting evidence of convergence for E10 and E11, while E12 reports a deterioration of that evidence since the 2008 financial crisis. In the absence of the restriction, evidence of convergence is sporadic after 2004 for E11 and almost disappears for E10 and E12. It should perhaps be emphasized that, for the particular case of testing for group-wise convergence, there is no question that imposing the restriction that the sum of the intercepts in Equation (3) is equal to zero is the correct procedure. Unlike the usual case of imposing restrictions, which may or may not be correct, this restriction is correct by construction.

---

[24]For each window of estimation, a new p-value is generated using the bootstrap procedure explained in Section 2.3.

[25] As an illustration, for $(N,T,\rho) = (10,100,0.97)$, the restricted test has a power of 0.6 while the unrestricted has a power of 0.38

[26] There are no rejections for the windows ending in 2001:9-2002:1, 2005:7-2005:11, and 2009:1-2009:5.

Figure 3 plots the values of ρ for the restricted model for E10, E11 and E12 from 1979:1-1986:12 to 2002:5-2010:4. In accord with our definition of (stochastic) convergence, variations of ρ can be interpreted as a measure of the strength of inflation group-wise convergence. As a result, a more persistent differential (higher value of ρ) would correspond to an weaker convergence among Euro area inflation rates as any shock would have a longer lasting impact, and vice versa.[27] Unlike the p-values, the rate of convergence observes three clear periods. First, it remains relatively stable up to the window ending in 1997:3 (ρ close to 0.96 for both panels, E10 and E11). Then, both panels report drastic changes: the windows ending between 2002:2 and 2003:12, first, report a significant reduction in persistence (ρ decreases from 0.939 to 0.839 for E10 and from 0.945 to 0.866 for E11), which is then partially compensated by a strengthening of the persistence (ρ increases from 0.842 to 0.898 for E10 and from 0.866 to 0.906 for E11) for the windows ending in 2004:1-2004:12. Following this period of transition, a period of stability concludes the sample: the windows ending between 2005:1 and 2010:04 report an average value for ρ of 0.898 for E10 and 0.900 for E11. In contrast, E12 observes relatively stable period up to the window ending in 2009:1, then the average ρ increases from 0.913 to 0.927 for the remaining of the sample.

Both E10 and E11 end with the strongest degree of convergence among the inflation rates of the entire period studied. The lower values of ρ for the windows starting in 2002:2 are consistent with Honohan and Lane's (2004) evidence of convergence in Euro area inflation rates since 2002.[28] In contrast, the E12 panel reports very different results showing a weaker degree of convergence among this group of inflation rates (higher ρ than for E10 or E11), weakened even more after the financial crisis.

The behavior of both the p-values and the rates of convergence for E10 and E11 closely follow the European Monetary Union (EMU) timetable up to the financial crisis. The mechanism that led to the single currency included three major steps: from 1990:7 to 1993:12 (windows ending in 1997:7-2000:12), capital was allowed to move freely within the European Economic Community, from 1994:4 to 1998:12 (windows ending in 2001:4-2005:12) the

---

[27]In other words, the time series definition of convergence is that the difference between the series is stationary. Hence, ρ, the rate of convergence, measures the strength of the relationship among these series.

[28] While the value of ρ is biased downward, the focus in the section is on a comparison across time periods which would not be affected by bias correction. In the next section, we conduct median-unbiased estimation for several windows.

Treaty of Maastricht was implemented and in 1999:1 (window ending in 2006:12), the single currency was introduced. Finally, the 2008 financial crisis is noticeable, especially for E12, confirming the impact of the Irish crisis on the group convergence.

Finally, evidence of group-wise stationarity for E10 and E11 occurs several years before the processes reach a steady level of persistence. While inflation rates have converged shortly after the implementation of the Maastricht treaty, they do not attain a stable level of convergence until a year prior to the fixed parity between the exchange rates. Even though the 2008 financial crisis led to a temporary increase in ρ, which reaches its highest point of 0.915 with the window ending in 2009:4 for E10 and of 0.913 for E11, the final degree of convergence is significantly higher (ρ is significantly lower) than the one estimated for the first two phases of the EMU ($\rho = 0.895$ and 0.886 for E10 and E11, respectively).

3.2 Behavior of Euro area inflation rates as a group: measuring persistence

A closer investigation of the impact of the Euro and of the financial crisis requires a more accurate measure of ρ. Hence, we apply median unbiased corrections to the restricted and the unrestricted estimates. We focus on four windows: 1982:7-1990:6, or the pre-Maastricht era, 1990:7-1998:6, or the pre-Euro period, which ends six months before the exchange rates were definitely fixed, and two Euro periods: 1999:1-2006:12, just after the advent of the Euro and 2002:5-2010:4, which includes for the 2008 financial crisis.

Following Murray and Papell (2005), we use an extension of the Andrew and Chen (1994) method to the panel framework. The originality of our approach, however, consists of generating median unbiased estimates of the homogeneous rate of convergence for the restricted model. The iterative procedure used to generate the approximately median unbiased estimate, ρ_{AMU}, of ρ in (3) starts with the estimation of ϕ_{ij} in (3) via the new procedure. Then, assuming the estimates of ϕ_{ij}'s are true, the first median unbiased estimate $\rho_{1,AMU}$ is obtained by finding the median-unbiased estimator that corresponds to the value of $\rho_{SUR-restricted}$. We then assume $\rho_{1,AMU}$ to be the true value of ρ and obtain a new set of estimates for the ϕ_{ij}'s. Conditional on these news estimates, we obtain the new median unbiased estimates $\rho_{2,AMU}$. The iterative process continues until convergence occurs and median unbiased estimates of $\rho_{SUR-restricted}$ and the ϕ_{ij}'s are obtained.

Table 3 reports the rates of convergence for the differentials, the median unbiased estimates (point estimates and 95% confidence intervals of ρ), and the corresponding half-lives. The median unbiased point estimates are (as expected) higher than the GLS estimates. The Euro periods are characterized by the fastest rates of convergence, followed by the pre-Maastricht period, with the pre-Euro period displaying the slowest convergence rates. This pattern holds for the E10 and E11 panels and the restricted and unrestricted estimates.[29] For example, using the restricted model, E10 demonstrates a strengthening in group-wise inflation convergence as ρ_{MU} decreases from 0.970 for the pre-Maastricht period and 0.975 for the pre-Euro to 0.940 for the Euro periods.

As expected, there is no difference between the restricted and unrestricted GLS estimates, because the restriction is respected by the data. However, the restriction leads to a smaller variance of the estimates which is confirmed by the lower restricted median-unbiased estimates across all periods. Similarly, all the confidence intervals when the restriction is imposed are narrower than the unrestricted confidence intervals, confirming the gain in precision from the restrictions discussed above.[30]

The 95% confidence intervals for the Euro periods confirm the stronger evidence of convergence among the Euro area inflations from the point estimates. The confidence intervals for the E10 panel with the restricted model widen between the pre-Maastricht (0.950 to 0.988) and the pre-Euro (0.946 to 0.996) periods. In contrast, the confidence intervals for the Euro period (0.905 to 0.973 and 0.899 to 0.972) have a smaller upper bound and a much smaller lower bound than the confidence intervals for the two earlier periods.

The persistence of an economic time series is commonly measured with the half-life, the number of periods it takes for a shock on the inflation differential to dissipate by 50 percent. The half-life is approximated by the ratio $(\ln(0.5)/\ln(\rho_{MU}))$.[31] The median unbiased estimates and corresponding confidence intervals for the half-lives provide a more explicit

---

[29] The only exception is for the unrestricted E10 panel, for which the value of ρ is slightly lower for the pre-Euro than for the pre-Maastricht period.

[30] While the E11 panel for the pre-Maastricht period appears to be an exception, with the width of the confidence interval equal to 0.68 for the restricted and 0.53 for the unrestricted estimates, that interpretation is not correct. The upper point of the confidence interval for the unrestricted model is 1.00. Since the confidence intervals are constrained not to exceed unity, no comparison can be made in this case.

[31] While it is generally preferable to compute half-lives from impulse response functions, the panel model used allows for different serial correlation across series, hence there is no common impulse response function on which the half life could be based.

illustration of the speed of convergence. A larger half-life would imply slower decay and weaker inflation group-wise convergence.

Our results once again illustrate the gain in information when using the restriction: the restricted HL_{MU} point estimates are consistently lower that unrestricted estimates. More importantly, the gain in precision leads to narrower restricted confidence intervals, with a noticeable difference for the upper boundaries. For the half-lives, every restricted confidence interval is narrower than the corresponding unrestricted confidence interval.

Since the restrictions are valid by construction, we will focus on the median-unbiased estimates of the restricted model. The half-lives of the point estimates for both E10 and E11 decrease by more than 40 percent between the pre-Maastricht and Euro periods and by more that 50 percent between the pre-Euro and Euro periods: E10 (E11)'s half-lives rose from 22.76 (23.55) months in the pre-Maastricht period to 27.38 (98.67) months in the pre-Euro period, followed by a decline to 11.40 (10.31) months in the first Euro period and to 11.20 (13.80) months after the 2008 crisis. The half-lives for the E10 and E11 panels are very similar for the pre-Maastricht and the first Euro period. They are, however, very different for the pre-Euro period with a drastic slowdown of the speed of convergence for E11 (increase in ρ_{MU}) after the Maastricht treaty, highlighting the impact of Greece and its difficulties in keeping its inflation under control. Similarly, E11 reports an increase in persistence in the last period, when compared to the previous period and to E10, which is coherent with Greece's recent crisis. The E12 panel exists only for the Euro periods; yet, its estimates confirm an overall more persistent behavior than E10 and E11.

An overall decrease in persistence and narrowing of HL_{MU}'s confidence intervals for the Euro periods is a robust result through both E10 and E11 panels. Going from the pre-Maastricht to the pre-Euro period, the confidence intervals of the half-lives widen for the E10 panel, due to the numerous changes Europe had in the early 1990s (German reunification, different economic policies) and its evolution toward the more rigorous structure defined by the Maastricht treaty. Similarly, for the same periods, the confidence intervals for the E11 panel increase and widen, again reflecting the influence of the inclusion of Greece. However, as previously reported for the point estimates, the HL_{MU}'s confidence interval of E11 has a higher upper bound when the financial crisis is included. The persistence measured remains

significantly lower than prior to the Euro, which is in line with the message of Figure 1: the inflations report very similar behavior except for Greece and Ireland.

3.3 The ECB's goal of price stability

Can we reconcile our results with the ECB's medium run target of a 2% inflation rate or lower for the Euro area? We address this question by first looking at Euro area inflation rates and then by isolating the countries with consistently higher rates.

The average Euro area inflation rates for E10, E11 and E12, reported in Figure 4a, are close to the ECB's definition of price stability since the advent of the Euro, with few temporary out-of-target periods: 2000-2002, 2007-2008 and end of 2009. It is worth noting that the period of price stability corresponds to the periods of group-wise convergence and of lower levels of persistence (i.e strong convergence among the inflation rates).

Our results, also, show that the inclusion of Greece and Ireland weakened this group-wise convergence, especially after the 2008 financial crisis. A closer look at each individual country's inflation rate provides some insights on the type of convergence existing within the 12 initial Euro countries. Clearly, the inflation rates can observe either a similar level of inflation or consistently different rates ($\eta_i = 0$ or $\neq 0$ in Equation 1, respectively), while still respecting the ECB's goal of price stability.

We set as a benchmark the group of the seven lowest inflation rates, E7, (Austria, Belgium, Finland, France, Germany, Luxembourg, and Netherland). Figure 4b plots the confidence intervals corresponding to the E7 cross-sectional mean plus/minus two times its standard deviation and the inflation rates of the five remaining countries.[32] As a basic rule of thumb, any inflation out of this area is quite different from the seven others. Until 2000, Greece, Ireland, Italy, Portugal, and Spain report the highest inflation rates. Between 2000 and 2008, Greece and Spain report inflation rates that are among the highest of the group while Ireland's rate is clearly much higher than the rest of the group. Such inflation differentials imply a loss in relative price competitiveness for the countries with higher than average inflation rates, leading to issues such as trade imbalances within the single currency area. Hence, it is not surprising that Greece, Ireland and Spain were the most affected by the 2008 crisis within the Euro area.

---

[32] Results qualitatively similar have been obtained by plotting the c_i when considering differential with respect to the E7 cross-sectional mean for the same 5 countries.

After the crisis, Ireland's inflation rate falls precipitously, causing real depreciation that may help toward a faster recovery, unlike Spain's inflation that observes a moderate decrease. Concurrently, Greece still reports a much higher inflation rate than the Euro area as a whole. Ireland and Greece offset each other impact on the cross-sectional mean of the group. As a result, the inflation of the Euro area (based on the initial 12 countries) remains within the ECB's benchmark.

4. Conclusions

This paper investigates the behavior of inflation rates within the Euro area from 1979:1 to 2010:4. More specifically, it focuses on the impact of events such as the advent of the Euro and the recent financial crisis on the countries' ability to follow ECB's price stability criterion. The analysis relies on rolling an 8-year window through the entire sample, from 1979:1-1986:12 to 2002:5-2010:4, isolating Euro periods that both do and do not include the 2008 crisis.

The main difficulty in focusing on such a short window of estimation is the poor performance of standard time series tools. We propose a new estimation procedure that can be used when investigating convergence of a group of series toward a common target. Group-wise time series convergence is commonly measured using panel unit root tests on differentials generated as the difference between each series and the cross-sectional average. Hence, each resulting differential has a non-zero mean, but the cross-sectional mean of the group of differentials is equal to zero for each period. Our method uses that information in order to increase the size adjusted power of the test. Monte Carlo simulations report noticeable improvements of the test's power, especially when the data is persistent data ($\rho >$ 0.9) or when the data has a limited length ($T < 200$). Both of these characteristics are commonly featured in macroeconomic time series. Furthermore, the restricted ADF-SUR test also shows a greater ability of rejecting the unit root null solely when all the series are stationary, which is a welcome improvement on one of the most acknowledged drawbacks of the panel unit root approach. This estimation procedure could be useful when assessing the groupwise convergence of other macroeconomic series such as output growth, worked hours, and total factor productivity.

The increase in size adjusted power from the imposition of the true restriction allows us to estimate the model for the pre-Maastricht, pre-Euro, and Euro periods. Our results show that, while sporadic evidence of inflation convergence begins shortly after the implementation of the Maastricht treaty; steady evidence that the series have converged only occurs during the Euro periods. The median-unbiased estimate of the rate of convergence is much faster and the corresponding confidence intervals are considerably narrower for the Euro periods than for the two earlier periods. The half-lives of the point estimates of the differentials, the number of periods that it takes for a shock to the inflation differentials to decrease by one-half, falls by more than 40 percent between the pre-Maastricht and Euro periods and by more that 50 percent between the pre-Euro and Euro periods.

The compelling evidence of group-wise convergence among the initial Euro area countries presented here for the Euro periods also leads to another important point: this same group of countries meets the ECB's objective on price stability. This result holds even when we show that a minority of countries, namely Greece, Ireland and Spain, report consistently higher inflation rates for most of the Euro period. The resulting loss in price competitiveness contributed to the trade imbalances and current account deficits that partly explain the buildup of bubbles and these countries' disadvantage in advance of the 2008 crisis. Interestingly, the shock of the financial crisis may have provided the impetus that forced these countries to control their inflations. This adjustment seems quite successful for Ireland which has an inflation rate below the group tendency after the 2008 crisis, and for Spain. Greece, however, remains an outlier as it remains unable to control its inflation at the end of the period.

References

Andrews, D. W. K., and H-Y. Chen, 1994, Approximately Median-Unbiased Estimation of Autoregressive Models, *Journal of Business & Economic Statistics*, 12 (2): 187-204.

Bai, J. and S. Ng, 2004, A PANIC Attack on Unit Roots and Cointegration, *Econometrica*, 72: 1127-1177.

Ben-David, D., 1993, Equalizing Exchange: Trade Liberalization and Income Convergence, *Quarterly Journal of Economics*, 108(3): 653-79.

Ben-David, D., 1996, Trade and convergence among countries, *Journal of International Economics*, 40(3-4): 279-298.

Bernard, A. B. and S.N. Durlauf, 1995, Convergence in international output, *Journal of Applied Econometric*, 10: 97-108.

Bernard, A. B. and S.N. Durlauf, 1996, Interpreting tests of the convergence hypothesis, *Journal of Econometrics*, 71: 161-174.

Breitung, J. and M. H., Peseran, 2005, Unit Roots and Cointegration in Panels, L. Matyas and P. Sevestre (eds), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*.

Breuer, J., R. Robert and M. Wallace, 2001, Misleading Inferences from Panel Unit-Root Tests with an Illustration from Purchasing Power Parity, *Review of International Economics*, 9(3): 482-93.

Breuer, J., R. McNown and M. Wallace, 2002, Series-Specific Unit Root Tests with Panel Data, *Oxford Bulletin of Economics and Statistics*, 64(5):527-46.

Busetti F., L. Forni, A. Harvey and F. Venditti, 2007, Inflation Convergence and Divergence within the European Monetary Union, *International Journal of Central Banking, International Journal of Central Banking*, 3(2): 95-121.

Chang, Y. 2004, Bootstrap Unit Root Tests in Panels with Cross-sectional Dependency, *Journal of Econometrics*, 120: 263-293.

Davidson, R., and J.G MacKinnon, 1999, The Size Distortion Of Bootstrap Tests. *Econometric Theory*, 15:361-376.

Davidson, R. and J.G. MacKinnon, 2006, Bootstrap methods in econometrics, *Palgrave Handbooks of Econometrics: Volume 1 Econometric Theory*, 812-838.

Durlauf, S. and Quah, 1999, The new empirics of economic growth, Handbook of Macroeconomics, in: J. B. Taylor & M. Woodford (ed.), Handbook of Macroeconomics, ed.1, 1 (4): 235-308.

European Central Bank, 2003, Inflation differentials in the euro area: potential causes and policy implications

Engel, C. and J. Rogers, 2004, European Market Integration After the Euro, *Economic Policy*, 19: 347-384.

Evans, P. and G. Karras, 1996, Do Economies Converge? Evidence from a Panel of U.S. States, *The Review of Economics and Statistics*, 78(3): 384-88.

Evans, P. 1998, Using Panel Data to Evaluate Growth Theories, *International Economic Review*, 39(2): 295-306.

Fleissig, A.R. and J. Strauss, 2001. Panel unit-root tests of OECD stochastic convergence, *Review of International Economics*, 9: 153–162.

Fritsche, U. and V. Kuzin, 2008, Analyzing Convergence in Europe Using a Non-linear Single Factor Model, Macroeconomics and Finance Series, Hamburg University.

Greene, W.H, 2008, Econometric Analysis, Pearson Hall, 6th edition.

Hall, A., 1994, Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection, *Journal of Business and Economic Statistics*, 12:461-470.

Honohan, P. and P.R Lane, 2003, Divergent inflation rates in EMU, Economic Policy, 12 (37): 357-394.

Honohan, P. and P.R Lane, 2004, Exchange Rates and Inflation Under EMU: An Update, CEPR Discussion Papers 4583.

Im, K.S., M.H. Pesaran, and Y., Shin, 2003, Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, 115(1): 53-74.

Islam, N., 1995, Growth Empirics: A Panel Data Approach**, *Quarterly Journal of Economics*, 110(4): 1127-70.

Judge, G.G., Hill, R.C., Griffiths, W., Lütkepohl, H. and T. Lee, 1988, Introduction to the Theory and Practice of Econometrics, Wiley, 2nd edition.

Kočenda, E. and D.H. Papell, 1997. Inflation Convergence Within the European Union: A Panel Data Analysis, *International Journal of Finance and Economics*, 2(3): 189-198.

Lee, H. Y. and J. L. Wu, 2001, Mean Reversion of Inflation Rates: Evidence from 13 OECD Countries, *Journal of Macroeconomics*, 23: 477-487.

Levin, A., C.F. Lin, and C.J. Chu, 2002, Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties, *Journal of Econometrics*, 108: 1-24.

Lopez, C., 2009a, Panel Unit Root with Good Power in Small Samples, *Econometric Reviews*, 28 (4): 295-313

Lopez, C, 2009b, Euro-zone Inflation Rates: Stationary or Regime-wise Stationary Processes, *Economics Bulletin*, 29(1): 238-243

Maddala, G.S. and S. WU, 1999, A comparative Study of Unit Root Tests with Panel Data and a New Simple Test, Oxford *Bulletin of Economics and Statistics*, 61: 631-652.

McKinnon, J., 2006, Bootstrap Methods in Econometrics, The Econometrics RecordThe Economic Society of Australia, 82(1): S2-S18.

Murray, C.J., and D.H. Papell, 2005, Do Panels Help Solve the Purchasing Power Parity Puzzle?, 2005, *Journal of Business and Economic Statistics,* 23: 410-415.

O'Connell P., 1998, The Overvaluation of Purchasing Power Parity, *Journal of International Economics*, 44: 1-19.

Papell, D.H., 1997, Searching for Stationarity: Purchasing Power Parity under the Current Float, *Journal of International Economics*, 43: 313-332.

Pesaran, M.H, 2007, A pair-wise approach to testing for output and growth convergence, *Journal of Econometrics*, 138(1): 312-355.

Rogers, J.H., 2007, Monetary Union, Price Level Convergence, and Inflation: How Close is Europe to the USA?, *Journal of Monetary Economics*, 54 (3): 785-796.

Rogers, J.H., Hufbauer, G., Wada, E., 2001. Price level convergence and inflation in Europe, Institute for International Economics, WP 01-1.

Sarno, L. and M.P. Taylor, 2002, Purchasing Power Parity and the Real Exchange Rate, *International Monetary Fund Staff Papers*, 49(1), 65-105.

Taylor, M.P, and L. Sarno, 1998, The behavior of real exchange rates during the post-Bretton Woods period, *Journal of International Economics*, 46 (2): 281-312.

Taylor, A.M. and M.P. Taylor, 2004, The Purchasing Power Parity Debate, *Journal of Economic Perspectives*, 18(4):135-158.

Weber A.A and G.W. Beck, 2005, Price Stability, Inflation Convergence and Diversity in EMU: Does One Size Fit All?, CFS WP 2005/30.

Table 1: Finite Sample Performance of the Restricted and Unrestricted ADF-SUR test

DGP: $y_{it} = \rho y_{i,t-1} + u_{i,t}$ with $i = 1,..., N$ $t = 1,..., T$ and $u_{it} \sim WN$

Estimated model: $\Delta y_{it}^{diff} = c_i + \rho y_{i,t-1}^{diff} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j}^{diff} + \varepsilon_{it}$

N	T	$\rho=1.00$		0.99		0.97		0.95		0.90	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
5	25	0.052	0.052	0.078	0.057	0.127	0.071	0.187	0.087	0.302	0.164
	50	0.051	0.053	0.089	0.067	0. 177	0.102	0.325	0134	0.800	0.455
	100	0.052	0.051	0.125	0.081	0.410	0.225	0.769	0.503	0.999	0.960
	200	0.051	0.052	0.187	0.144	0.865	0.679	1.000	0.988	1.000	1.000
10	25	0.051	0.058	0.081	0.061	0.136	0.073	0.202	0.094	0.435	0.184
	50	0.050	0.052	0.102	0.074	0.248	0.138	0.394	0.267	0.918	0.764
	100	0.050	0.052	0.155	0.112	0.595	0.384	0.942	0.827	1.000	1.000
	200	0.050	0.049	0.357	0.223	0.993	0.947	1.000	1.000	1.000	1.000
20	25	0.049	0.048	0.053	0.049	0.057	0.051	0.062	0.050	0.074	0.054
	50	0.050	0.048	0.106	0.078	0.290	0.167	0.539	0.337	1.000	0.918
	100	0.053	0.053	0.228	0.166	0.814	0.664	1.000	0.990	1.000	1.000
	200	0.051	0.051	0.591	0.373	1.000	0.99	1.000	1.000	1.000	1.000

(1) corresponds to the restricted model that uses $\sum_{i=1}^{N} c_i = 0$, while (2) stands for the unrestricted case. Reading illustration: if $(N, T, \rho) = (5, 100, 0.97)$, the size adjusted power of the restricted ADF-SUR test is of 0.410 compared to 0.225 for the unrestricted case.

Table 2: Finite Sample Power of the Restricted and Unrestricted ADF-SUR test

Mixed processes, T=100

DGP: $y_{it} = \rho_i y_{i,t-1} + u_{i,t}$ with $u_{it} \sim WN$

Where $\rho_i < 1$ for $i = 1,...,k$ and $\rho_j = 1$ for $j = k+1,...,N$.

Estimated model: $\Delta y_{it}^{diff} = c_i + \rho y_{i,t-1}^{diff} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j}^{diff} + \varepsilon_{it}$

N				5				
ρ_i	0.97		0.95		0.90		0.80	
$\rho_j = 1$	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
N-k								
0	0.41	0.23	0.77	0.50	1.00	0.96	1.00	1.00
1	0.22	0.16	0.37	0.31	0.58	0.65	0.70	0.81
2	0.15	0.12	0.21	0.19	0.31	0.36	0.40	0.51
3	0.09	0.09	0.12	0.12	0.16	0.19	0.22	0.26
4	0.07	0.07	0.08	0.08	0.09	0.10	0.11	0.11
5	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

(1) corresponds to the restricted model that uses $\sum_{i=1}^{N} c_i = 0$, while (2) stands for the unrestricted case.

Reading illustration: if N-k=2 then the panel is a mix of 2 unit roots and 3 stationary processes. If then ρ_i =0.97, the size adjusted power of the restricted ADF-SUR test is of 0.15 compared to 0.12 for the unrestricted case.

Table 2 (continue): Finite Sample Power of the Restricted and Unrestricted ADF-SUR test

Mixed processes, T=100

DGP: $y_{it} = \rho_i y_{i,t-1} + u_{i,t}$ with $u_{it} \sim WN$

Where $\rho_i < 1$ for $i = 1,...,k$ and $\rho_i = \rho_j = 1$ for $j = k+1,...,N$.

Estimated model: $\Delta y_{it}^{diff} = c_i + \rho y_{i,t-1}^{diff} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j}^{diff} + \varepsilon_{it}$

N	10							
ρ_i	0.97		0.95		0.90		0.80	
$\rho_j = 1$	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
N-k								
0	0.60	0.38	0.94	0.83	1.00	1.00	1.00	1.00
1	0.43	0.33	0.71	0.65	0.92	0.93	0.97	0.98
2	0.32	0.26	0.51	0.50	0.74	0.82	0.86	0.92
3	0.25	0.21	0.38	0.38	0.57	0.66	0.70	0.80
4	0.19	0.17	0.28	0.29	0.42	0.49	0.54	0.64
5	0.15	0.14	0.21	0.22	0.30	0.36	0.39	0.48
6	0.12	0.11	0.15	0.16	0.21	0.25	0.27	0.33
7	0.09	0.09	0.11	0.12	0.15	0.17	0.18	0.22
8	0.08	0.07	0.09	0.09	0.10	0.11	0.12	0.13
9	0.06	0.06	0.07	0.07	0.07	0.08	0.08	0.08
10	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

(1) corresponds to the restricted model that uses $\sum_{i=1}^{N} c_i = 0$, while (2) stands for the unrestricted case.

Reading illustration: if N-k=4 then the panel is a mix of 4 unit roots and 6 stationary processes. If then ρ_i =0.90, the size adjusted power of the restricted ADF-SUR test is of 0.42 compared to 0.49 for the unrestricted case.

Table 2 (continue): Finite Sample Power of the Restricted and Unrestricted ADF-SUR test, Mixed processes, T=100

DGP: $y_{it} = \rho_i y_{i,t-1} + u_{i,t}$ with $u_{it} \sim WN$

Where $\rho_i < 1$ for $i = 1,...,k$ and $\rho_i = \rho_j = 1$ for $j = k+1,...,N$.

Estimated model: $\Delta y_{it}^{diff} = c_i + \rho y_{i,t-1}^{diff} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j}^{diff} + \varepsilon_{it}$

N	20							
ρ_i	0.97		0.95		0.9		0.8	
$\rho_j=1$	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
N-k								
0	0.81	0.66	1.00	0.99	1.00	1.00	1.00	1.00
1	0.69	0.58	0.95	0.93	0.99	1.00	1.00	1.00
2	0.60	0.52	0.87	0.86	0.98	0.99	0.99	1.00
3	0.52	0.46	0.79	0.79	0.94	0.97	0.98	1.00
4	0.39	0.40	0.69	0.71	0.88	0.94	0.95	0.99
5	0.33	0.36	0.60	0.63	0.80	0.89	0.90	0.97
6	0.29	0.31	0.52	0.54	0.72	0.82	0.84	0.93
7	0.25	0.27	0.44	0.47	0.62	0.74	0.76	0.88
8	0.23	0.24	0.37	0.40	0.53	0.65	0.67	0.81
9	0.21	0.21	0.27	0.34	0.46	0.56	0.58	0.71
10	0.19	0.18	0.25	0.29	0.39	0.48	0.50	0.63
11	0.16	0.16	0.23	0.25	0.32	0.40	0.42	0.53
12	0.14	0.14	0.19	0.20	0.26	0.33	0.35	0.44
13	0.11	0.12	0.16	0.17	0.22	0.26	0.28	0.35
14	0.10	0.11	0.14	0.15	0.18	0.21	0.23	0.27
15	0.09	0.10	0.11	0.12	0.14	0.16	0.18	0.21
16	0.08	0.08	0.09	0.10	0.11	0.13	0.14	0.16
17	0.07	0.07	0.08	0.08	0.09	0.10	0.11	0.12
18	0.06	0.06	0.07	0.07	0.08	0.08	0.08	0.09
19	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.07
20	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

(1) corresponds to the restricted model that uses $\sum_{i=1}^{N} c_i = 0$, while (2) stands for the unrestricted case.

Reading illustration: if N-k=8 then the panel is a mix of 8 unit roots and 12 stationary processes. If then ρ_i =0.90, the size adjusted power of the restricted ADF-SUR test is of 0.53 compared to 0.65 for the unrestricted case.
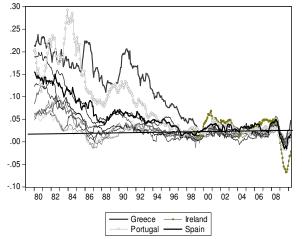
Table 3: Persistence Measurement: Median Unbiased Estimator (ρ_{MU}) and Half-Life (HL$_{MU}$=ln (0.5)/ln(ρ_{MU}))
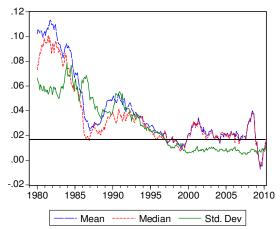
$$\Delta y_{it}^{diff} = c_i + \rho y_{i,t-1}^{diff} + \sum_{j=1}^{k_i} \phi_{ij} \Delta y_{i,t-j}^{diff} + \varepsilon_{it} \text{ , with } i = 1,...,N \quad t = 1,...,T \text{ and } \varepsilon_{it} \sim WN$$

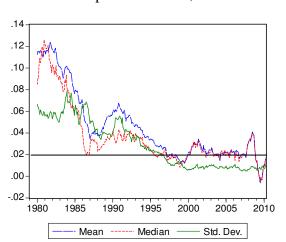|	ρ	ρ_{MU}	95%CI	HL$_{MU}$	95%CI
Restricted ADF-SUR estimation using			$\sum_{i=1}^{N} c_i = 0$		
E10					
1982:7-1990:6	0.958	0.970	(0.950; 0.988)	22.76	(13.51; 57.41)
1990:7-1998:6	0.943	0.975	(0.946; 0.996)	27.38	(12.48; 172.94)
1999:1-2006:12	0.895	0.941	(0.905; 0.973)	11.40	(6.94; 25.32)
2002:5-2010:4	0.893	0.940	(0.899; 0.972)	11.20	(6.51; 24.41)
E11					
1982:7-1990:6	0.952	0.971	(0.921, 0.989)	23.55	(8.42; 62.67)
1990:7-1998:6	0.957	0.993	(0.965; 0.999)	98.67	(19.46; 692.80)
1999:1-2006:12	0.895	0.935	(0.905; 0.974)	10.31	(6.94; 26.31)
2002:5-2010:4	0.910	0.951	(0.919; 0.979)	13.80	(8.21; 32.66)
E12					
1999:1-2006:12	0.915	0.956	(0.917; 0.983)	15.40	(8.00; 40.42)
2002:5-2010:4	0.929	0.960	(0.935; 0.985)	16.98	(10.31; 45.86)
Unrestricted ADF-SUR estimation					
E10					
1982:7-1990:6	0.957	0.979	(0.955; 0.998)	32.66	(15.05; 346.23)
1990:7-1998:6	0.942	0.977	(0.944; 1.000)	29.79	(12.03; ∞)
1999:1-2006:12	0.895	0.947	(0.903; 0.987)	12.73	(6.79; 52.97)
2002:5-2010:4	0.893	0.945	(0.884; 0.989)	12.25	(5.62; 62.67)
E11					
1982:7-1990:6	0.952	0.973	(0.947; 1.000)	25.32	(12.72; ∞)
1990:7-1998:6	0.955	0.994	(0.961; 1.000)	172.94	(17.42; ∞)
1999:1-2006:12	0.895	0.952	(0.911; 0.982)	14.09	(7.44; 38.16)
2002:5-2010:4	0.910	0.960	(0.916; 0.987)	16.98	(7.90; 52.97)
E12					
1999:1-2006:12	0.915	0.963	(0.928; 0.990)	18.38	(9.27; 68.97)
2002:5-2010:4	0.929	0.974	(0.942; 0.995)	26.31	(11.60; 138.28)

Fig. 1: Inflation Rates, Cross-sectional Mean, Median and Standard Deviation

Inflation rates 1979:1-2010:4

Descriptive Statistics, E10

Descriptive Statistics, E11
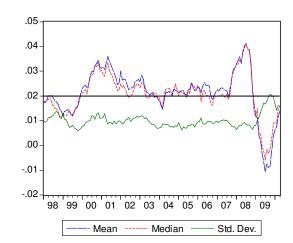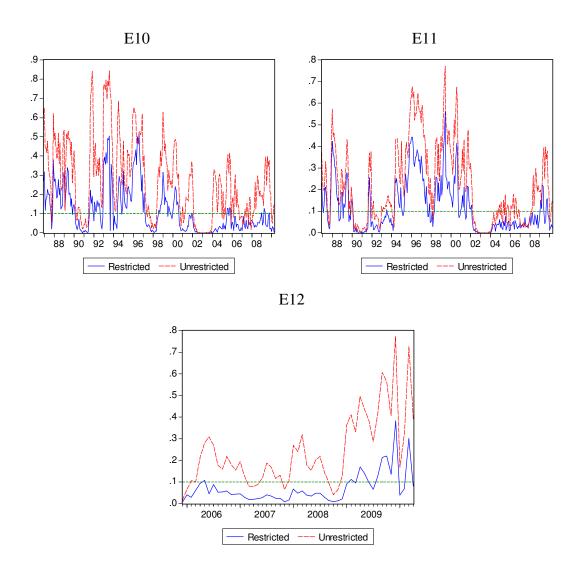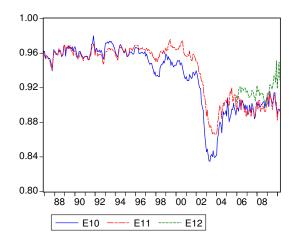

Descriptive Statistics, E12

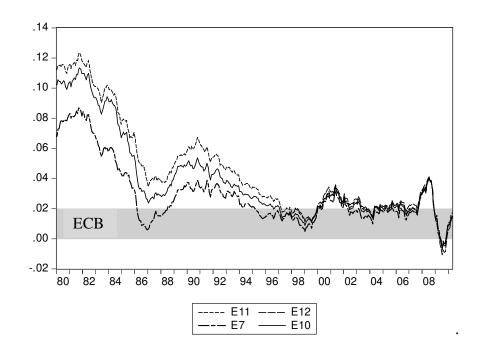Figure 2: P-values, rolling window from 1979:1-1986:12 to 2002:5-2010:4 [33]

E10

E11

E12

[33] The p-values are bootstrapped using the methodology explained in Section 2.3

Fig. 3: Homogeneous Rate of Convergence, rolling window from 1979:1-86:12 to 2002:5-10:4
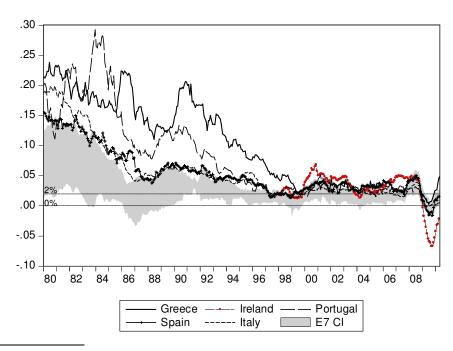

The x-axes report the end of the period estimated.

Fig. 4: Group and individual deviations from the ECB rule[34]

a. Group inflation rates

b. Individual inflation rates

[34] CI stands for central confidence interval based on two times the cross-sectional standard deviation of the group considered, E7.