COINTEGRATION TESTS OF PPP:

DO THEY ALSO EXHIBIT ERRATIC BEHAVIOUR?

Guglielmo Maria Caporale
Brunel University, London

Christoph Hanck
Universität Dortmund

September 2006

Abstract

We analyse whether tests of PPP exhibit erratic behaviour (as previously reported by Caporale et al., 2003) even when (possibly unwarranted) homogeneity and proportionality restrictions are not imposed, and trivariate cointegration (stage-three) tests between the nominal exchange rate, domestic and foreign price levels are carried out (instead of stationarity tests on the real exchange rate, as in stage-two tests). We examine the US dollar real exchange rate vis-à-vis 21 other currencies over a period of more than a century, and find that stage-three tests produce similar results to those for stage-two tests, namely the former also behave erratically. This confirms that neither of these traditional approaches to testing for PPP can solve the issue of PPP.

1   INTRODUCTION

Purchasing Power Parity (PPP) is one of the most popular theory for explaining the long-run behaviour of exchange rates, and has therefore been extensively investigated. Froot and Rogoff (1995) distinguish three stages in the time series literature on PPP. Stage-one tests were flawed by their failure to take into account possible non-stationarities in the series of interest. Stage-two tests focused on the null that the real exchange rate follows a random walk, the alternative being that PPP holds in the long run. However, such unit root tests were found to have very low power, and not to be able to distinguish between random-walk behaviour and very slow mean-reversion in the PPP-consistent level of the real exchange rate (see, e.g., Frankel, 1986, and Lothian and Taylor, 1997), unless very long spans of data were used (see, e.g., Lothian and Taylor, 1996, and Cheung and Lai, 1994). Stage-three tests have used cointegration methods, but essentially suffer from the same problem of low power, and consequently have not significantly improved our understanding of real exchange rate behaviour (see Rogoff, 1996).

Caporale et al. (2003) aimed to find an explanation for the contradictory evidence on PPP, even when long runs of data are used to increase the power of test statistics. They focused on stage-two tests and argued that the reason is that the type of stationarity exhibited by the real exchange rate cannot be accommodated by the fixed-parameter autoregressive homoscedastic models normally employed in the literature. Using a dataset including 39 countries and spanning a period of up to two centuries, they analysed the behaviour of both WPI- and CPI-based measures of the real exchange rate. In particular, they computed a recursive t-statistic, and showed that it has an erratic behaviour, suggesting the presence of endemic instability, and of a type of non-stationarity more complex than the unit root one usually assumed.

In the present study we explore this issue further by analysing whether erratic behaviour also characterises stage-three tests. The advantage of such tests is that they do not impose the homogeneity and proportionality restrictions entailed by stage-two tests, which might not hold in practice. Therefore, by carrying out cointegration tests of PPP we check whether there might be a relation between the presence of erratic

behaviour and the imposition of overly strong restrictions. The layout of the paper is as follows. Section 2 reviews the PPP condition in its different forms. Section 3 describes the data and presents some empirical evidence based on two different cointegration methods. Section 4 summarises the main findings and offers some concluding remarks.

2   THE PPP CONDITION

In its absolute form, the PPP condition states that the nominal exchange rate should be proportional to the ratio of the domestic to the foreign price level, i.e.:

$$s_t = \alpha + \beta_0 p_t - \beta_1 p_t^* \tag{1}$$

where s_t is the nominal exchange rate, p_t the domestic price level, and p_t^* the foreign price level, all in logs.[1] This is known as a trivariate relationship. Imposing the "symmetry" restriction $\beta_0 = -\beta_1 = \beta$ on the price coefficients, one obtains the following bivariate relationship:

$$s_t = \alpha + \beta(p_t - p_t^*) \tag{2}$$

Finally, the "proportionality" restriction $\alpha = 0, \beta = 1$ implies

$$q_t = s_t - p_t + p_t^* \tag{3}$$

where q_t is the real exchange rate.

Most of the literature in the 1980s tested PPP by means of (stage-two) unit root tests (DF or ADF – see Dickey and Fuller, 1979) on the real exchange rate, which, under PPP, should be stationary and revert to its long-run equilibrium value given by PPP after being hit by shocks. The null hypothesis is that it follows a random walk (it has a unit root), since market efficiency implies that its changes should be unpredictable, whilst the alternative is that PPP holds. The maintained (joint) hypothesis is that the symmetry/proportionality restrictions both hold, which might not be true in practice.

---

[1] Relative PPP implies that the percentage change in the exchange rate between two currencies equals the inflation differential, i.e. $\Delta s_t = \beta_0 \Delta p_t - \beta_1 \Delta p_t^*$.

Consequently, the evidence presented by Caporale et al. (2003) on the erratic behaviour of unit root tests might reflect unwarranted restrictions.

By contrast, a (stage-three) trivariate cointegration test of PPP entails running the following cointegrating regression (which does not impose any such restrictions):

$$s_t = \alpha + \beta_0 p_t - \beta_1 p_t^* + u_t \tag{4}$$

where the variables are defined as before, and u_t stands for the regression errors. PPP is then tested by means of DF and ADF tests on the estimated residuals. In the present paper, by implementing cointegration tests of this type, we aim to establish whether or not evidence of erratic behaviour can still be found, even without the abovementioned restrictions, and consequently whether or not the findings of Caporale et al. (2003) are robust or instead are due the imposition of unwarranted restrictions.

3   COINTEGRATION TESTS OF PPP

3.1 Data sources and definitions

We use the dataset also employed by Taylor (2002), which includes annual data for the nominal exchange rate, CPI and the GDP deflator. This dataset is particularly useful for our purposes because it covers a long period, ranging from 1892 through to 1996. The countries contained in our panel are given in Table 1. We use the United States as the reference country throughout. See Taylor (2002) for further details on data sources and definitions.

3.2 Empirical analysis

As a first step, we carried out standard augmented Dickey-Fuller (Said and Dickey, 1984) unit root tests to establish whether the series are all I(1), and it is therefore legitimate to test for cointegration. The results indicate that this is indeed the case (see Table 1). We then proceeded to the estimation of cointegrating regressions using the Engle and Granger (1987) methodology. That is, we estimated (4) recursively by OLS, and used the residuals to test the null hypothesis that they are nonstationary (i.e., that PPP does not hold) by means of DF and ADF tests. In order to investigate

possible parameter instability, we created a new time series "t-stat" which is the computed t-statistic from the recursive estimation of the coefficients of the following model whose order is selected using the Modified AIC (MAIC) of Ng and Perron (2001):

$$\Delta \hat{u}_t = \alpha_0 + \alpha_1 \hat{u}_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta \hat{u}_{t-j} + \varepsilon_t . \qquad (5)$$

Here, $\hat{u}_t$ are the residuals from OLS estimation of the cointegrating regression (4), ε_t is a white noise error term, and t-stat is defined as $\hat{\alpha}_1 / est.s.e.(\hat{\alpha}_1)$. Equation (4) is estimated recursively, using the first k observations to produce the first residual series, from which we compute the unit root test statistic $\hat{\alpha}_1 / est.s.e.(\hat{\alpha}_1)$. We then add an extra observation to compute the second estimate based on $k+1$ data points, and repeat the process until all T available observations have been used to yield $T - k + 1$ estimates of the test statistics. We let $k \approx 20 - 25$ to discard estimates which are heavily affected by small-sample bias. One can then plot the t-stat based on the recursive estimates to see more clearly whether it changes substantially as more data points are added, which would be a strong indication of instability in the parameter. Big jumps in either the rejection or the acceptance region, or from one to the other, are a strong sign of a structural break in the DGP.

The results are summarised in Table 2. Columns 4 and 5 show that the test decision on whether PPP holds or not is not constant over the sample in the vast majority of countries. Frequent switches from the rejection to the non-rejection regions are found to occur, the recursive t-statistic exhibiting erratic behaviour very similarly to the case of stage-two tests. For some graphical illustration, consider the cases of Argentina (Figure 1), Finland (Figure 2), Mexico (Figure 3), or Chile (Figure 4).[2] The instability found clearly does not concern specific points in time, such that it could be dealt with using procedures for cointegration testing in the presence of structural breaks (see, e.g., Hansen, 1992, or Gregory and Hansen, 1996), but appears instead to be of an endemic type. As a counterexample where no switches occur at the finite sample 5% level, see Denmark (Figure 5).

---

[2] The two lines at the bottom are the 10% and 5% critical values calculated as in MacKinnon (1991).

We conducted the same type of analysis using the GDP deflator this time to construct the real exchange rate, obtaining a very similar picture, namely erratic behaviour in the majority of cases. For instance, compare Figure 7 with the corresponding CPI based Figure 1. There are only a few exceptions, such as Denmark, where no rejections occur (Figure 8). [3]

Further, as a robustness check, we tried different number of lags in the ADF regressions (5). Overall, a qualitatively similar pattern emerges throughout, although we find that higher number of lags are associated with fewer rejections (see Figure 1 and Figure 9 to Figure 12). This is what one would expect, the estimation of too many parameters resulting in lower power (Phillips and Perron, 1988).

To explore more in depth the issue of possible structural breaks, we also used fixed-size windows.[4] That is, we select a fixed sample size T^* and create the nth entry of the series t-stat as before but now based on observations $t = n,...,T^* + n$, where $n = k,...,T - T^*$. One would expect using fixed windows to reduce the likelihood of structural breaks occurring within the chosen sample, and hence to result in more frequent rejections of the null hypothesis that PPP does not hold. However, it turns out that the behaviour of the t-stat series is, if anything, even more erratic for increasing window sizes. It appears that the answer to whether or not PPP holds is highly dependent on the chosen sample. For instance, using Danish data ending in the 1960s and early 70s an investigator using $T^* = 30$ years of data would strongly reject the null of PPP not holding (see Figure 6).

Finally, we carried out alternative cointegration tests in all cases. Specifically, we used the λ trace test (Johansen 1988, 1991). Here the critical values were obtained by modifying the asymptotic ones from Osterwald-Lenum (1992) using the response surface regression results of Cheung and Lai (1993). Some results are reported in Table 3.[5] Since this test statistic's null distribution is related to the χ^2 distribution,

---

[3] Results for other countries are available upon request.
[4] A variety of other methods could also be used to shed additional light on whether structural breaks are present (see, e.g., Ploberger and Krämer, 1996).
[5] Again, using WPI data or a different number of lagged differences in the Johansen procedure does not make a qualitative difference. Detailed results are available upon request.

unlike in the previous cases, the rejection region is now above the critical value lines. As can be seen, we find further evidence of erratic behaviour (Figure 13), suggesting that this is not due to the type of cointegration test used, but it is a more fundamental issue pertaining to the stochastic properties of the PPP relationship. Interestingly, switches from the rejection to the non-rejection region occur around the same time in a number of cases - compare, e.g., Figure 1 with Figure 13.[6]

4   CONCLUSIONS

In this paper we have analysed whether tests of PPP exhibit erratic behaviour (as previously reported by Caporale et al., 2003) even when (possibly unwarranted) homogeneity and proportionality restrictions are not imposed, and trivariate cointegration (stage-three) tests between the nominal exchange rate, domestic and foreign price levels are carried out (instead of stationarity tests on the real exchange rate, as in stage-two tests). We examine the US dollar real exchange rate vis-à-vis 21 other currencies over a period of more than a century, and find that stage-three tests produce similar results to those for stage-two tests, namely the former also behave erratically. This corroborates the findings of Caporale et al. (2003), in the sense that these do not appear to be the consequence of arbitrarily imposed (symmetry/proportionality) restrictions. [7]

Our results confirm that neither of the two traditional approaches to testing for PPP (stage-two and stage-three tests) can solve the issue of PPP. Consistently with Caporale et al. (2003), the reported evidence again points to some form of non-stationarity in the data which is unlike the standard unit-root type normally assumed, or even the "separable" type discussed in Caporale and Pittis (2002), but rather one

---

[6] Similar patterns emerge for Australia, Brazil, Canada, Denmark, Finland, France, Mexico, the Netherlands, Norway, Portugal, Sweden and the UK, that is 14 out of 19 countries for which the sample size is sufficiently large to make statistically meaningful statements.
[7] More recently, panel cointegration methods have been used to deal with the issue of the low power of time series tests of PPP (see, e.g., Pedroni, 2004, and also, for an extensive survey of the literature, Caporale and Cerrato, 2006). We are currently investigating whether such methods also produce erratic behaviour.

where all the unconditional moments are unknown functions of time. Future research should aim to determine its exact dynamic features. [8]

---

[8] Possible nonlinearities in exchange rates have increasingly become the focus of attention (see, e.g., Taylor, 2003.)

References

Caporale, G.M. and N. Pittis (2002), "Unit roots versus other types of time-heterogeneity, parameter time dependence and superexogeneity", *Journal of Forecasting*, 21,3, 207-223.

Caporale, G.M., N. Pittis and P. Sakellis (2003), "Testing for PPP: the erratic behaviour of unit root tests", *Economics Letters*, 80, 2, 277-284.

Caporale, G.M. and M. Cerrato (2006), "Panel data tests of PPP: a critical overview", *Applied Financial Economics*, 16, 1-2, 73-91.

Cheung, Y.-W. and K.S. Lai (1993), "Finite-sample sizes of Johansen's likelihood ratio tests for cointegration", *Oxford Bulletin of Economics and Statistics*, 55, 3, 313-328.

Cheung, Y.-W. and K.S. Lai (1994), "Mean reversion in real exchange rates", *Economics Letters*, 46, 3, 251-256.

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

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and error correction: representation, estimation and testing", *Econometrica*, 55, 251-276.

Frankel, J. (1986), "International capital mobility and crowding out in the U.S. economy: imperfect integration of financial markets or goods markets?" in R. Hafer (ed.), *How Open is the U.S. Economy?*, Lexington Books, Lexington.

Froot, K.A. and K. Rogoff (1995), "Perspectives on PPP and long-run real exchange rates", in G. Grossman and K. Rogoff (eds.), *The Handbook of International Economics*, vol. 3, Elsevier Press, Amsterdam.

Gregory, A. and B. E. Hansen (1996) "Residual–based tests for cointegration in models with regime shifts", *Journal of Econometrics*, 70, 99-126

Hansen B. E. (1992) "Tests for parameter instability in regressions with I(1) processes", *Journal of Business and Economic Statistics*, 10, 321-335

Johansen, S. (1988), "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, 12, 231-254

Johansen, S. (1991), "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models", *Econometrica*, 59, 1551-1580.

Lothian, J.R. and M. P. Taylor (1996), "Real exchange rate behavior: the recent float from the perspective of the past two centuries", *Journal of Political Economy*, 104, 3, 488-509.

Lothian, J.R. and M. P. Taylor (1997), "Real exchange rate behavior", *Journal of International Money and Finance*, 16, 6, 945-954.

MacKinnon, J.G. (1991), "Critical values for cointegration tests", in R.F. Engle and C.W.J. Granger (eds.), *Long-Run Economic Relationships*, 267-276, Oxford University Press, Oxford.

Ng, S. and Perron, P. (2001), "Lag length selection and the construction of unit root tests with good size and power", *Econometrica*, 69, 6, 1519-1554.

Osterwald-Lenum, M. (1992), "A note with quantiles of the asymptotic distribution of the ML cointegration rank statistics", *Oxford Bulletin of Economics and Statistics*, 54, 461-472.

Pedroni, P. (2004), "Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis", *Econometric Theory*, 20, 3, 597-625.

Phillips, P.C.B. and Perron, P. (1988), "Testing for a unit root in time series regression", *Biometrika*, 75, 2, 335-346.

Ploberger, W. and W. Krämer (1996), "A trend-resistant test for structural change based on OLS residuals", *Journal of Econometrics*, 70, 175-185.

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

Said, S.E. and Dickey, D.A. (1984), "Testing for unit roots in autoregressive-moving average models of unknown order", *Biometrika*, 71, 3, 599-607.

Taylor, A. (2002), "A century of purchasing-power parity", *Review of Economics and Statistics*, 84, 1, 139-150.

Taylor, M.P. (2003), "Purchasing Power Parity", *Review of International Economics*, 11, 3, 436-452.

Table 1: ADF Unit Root Tests

country	p_{CPI}	p_{GDP}	e
ARGENTINA	1.836	3.976	1.319
AUSTRALIA	-0.671	-0.906	0.578
BELGIUM	-1.666	-2.51	-0.771
BRAZIL	0.681	5.162	1.204
CANADA	-1.279	-2.079	0.875
CHILE	0.111	0.061	0.229
DENMARK	-1.941	-2.381	0.291
FINLAND	-1.158	-0.973	-0.763
FRANCE	-0.956	-0.849	-0.245
GERMANY	-2.19	-2.123	-1.525
ITALY	-0.843	-0.528	-0.527
JAPAN	-0.282	-1.189	-1.401
MEXICO	1.277	---[2]	1.751
NETHERLANDS	-1.875	-1.484	0.175
NEW ZEALAND	-0.953	---[2]	-0.313
NORWAY	-1.931	-2.188	0.017
PORTUGAL	-1.069	-1.089	-0.914
SPAIN	-0.314	-0.406	0.95
SWEDEN	-1.487	-2.226	0.185
SWITZERLAND	0.096	-0.526	0.151
UK	-0.472	-0.564	0.793
UNITED STATES	0.741		---[1]

N.A.: [1] reference country [2] series unavailable/too short

N.B. The number of lagged differences is chosen according to the MAIC (Ng and Perron, 2001). Yearly data from 1892 to 1996. p_{CPI} **is the log CPI price level,** p_{GDP} **is the log GDP deflated price level and** *e* **is the log nominal exchange rate.**
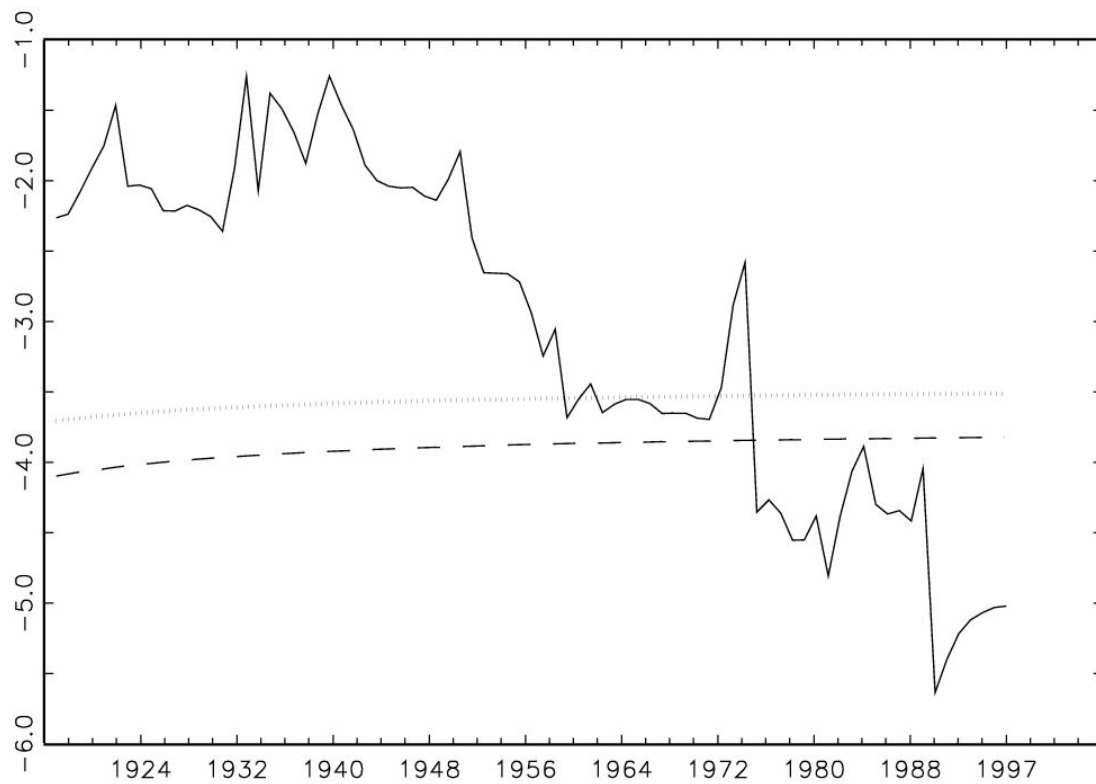
Figure 1: CPI-based Argentine t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression
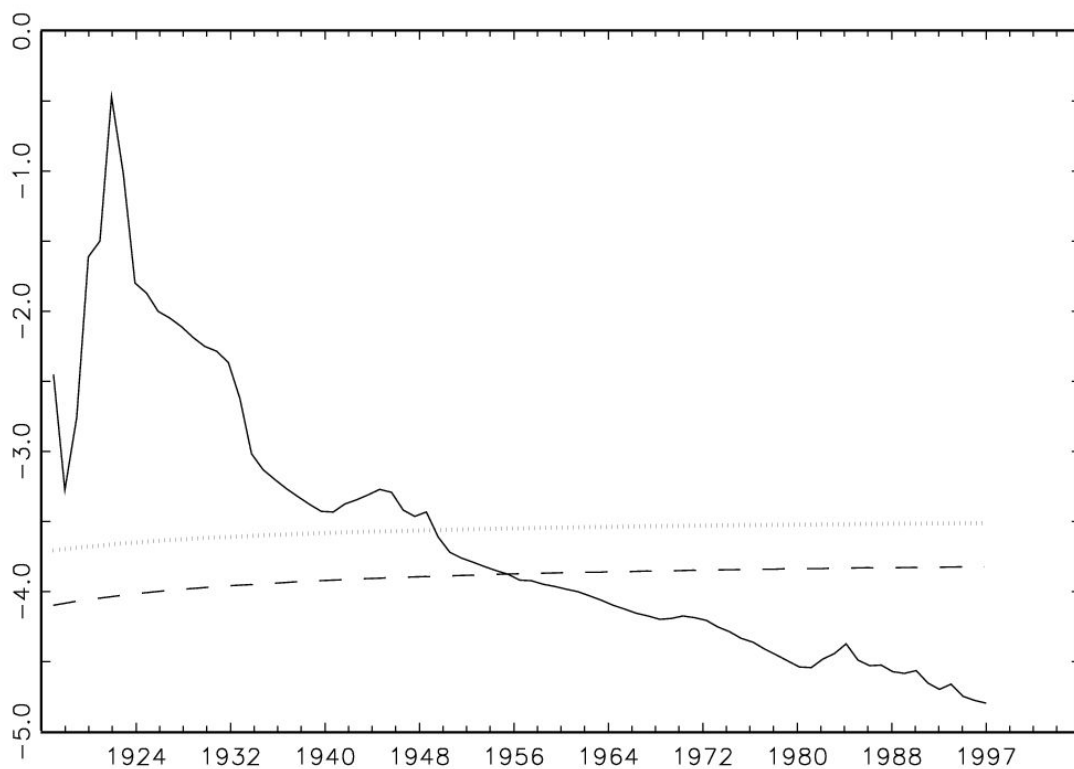

Figure 2: CPI-based Finnish t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression
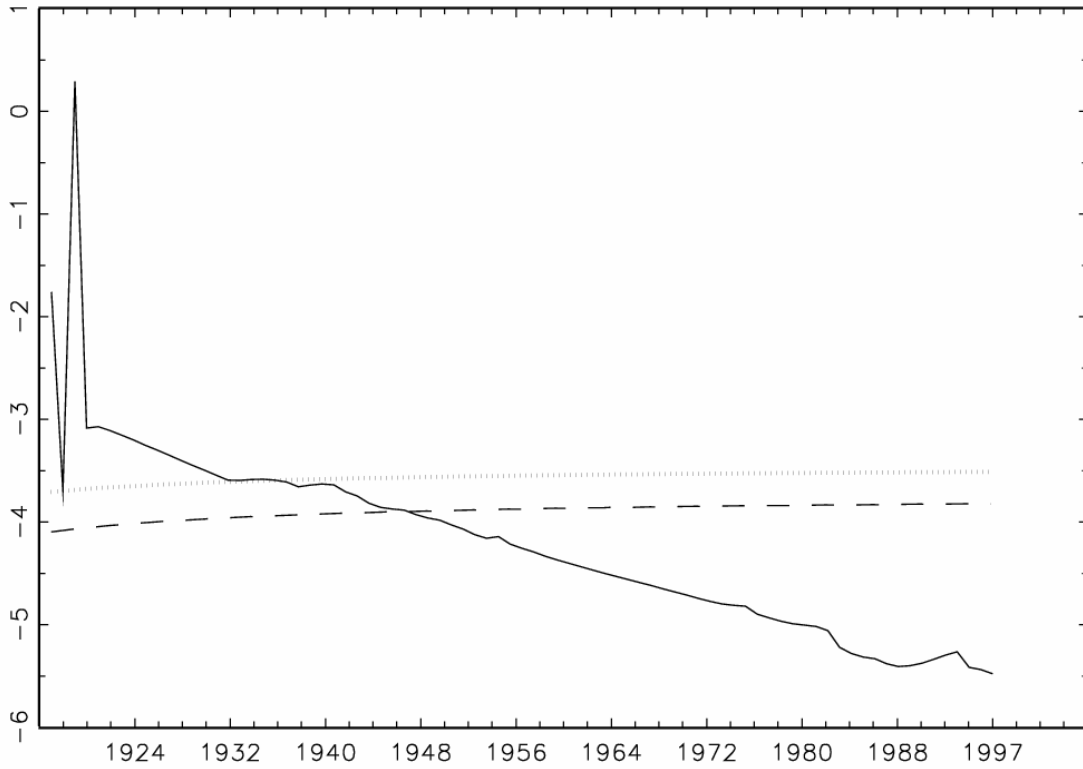
Figure 3: CPI-based Mexican t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression
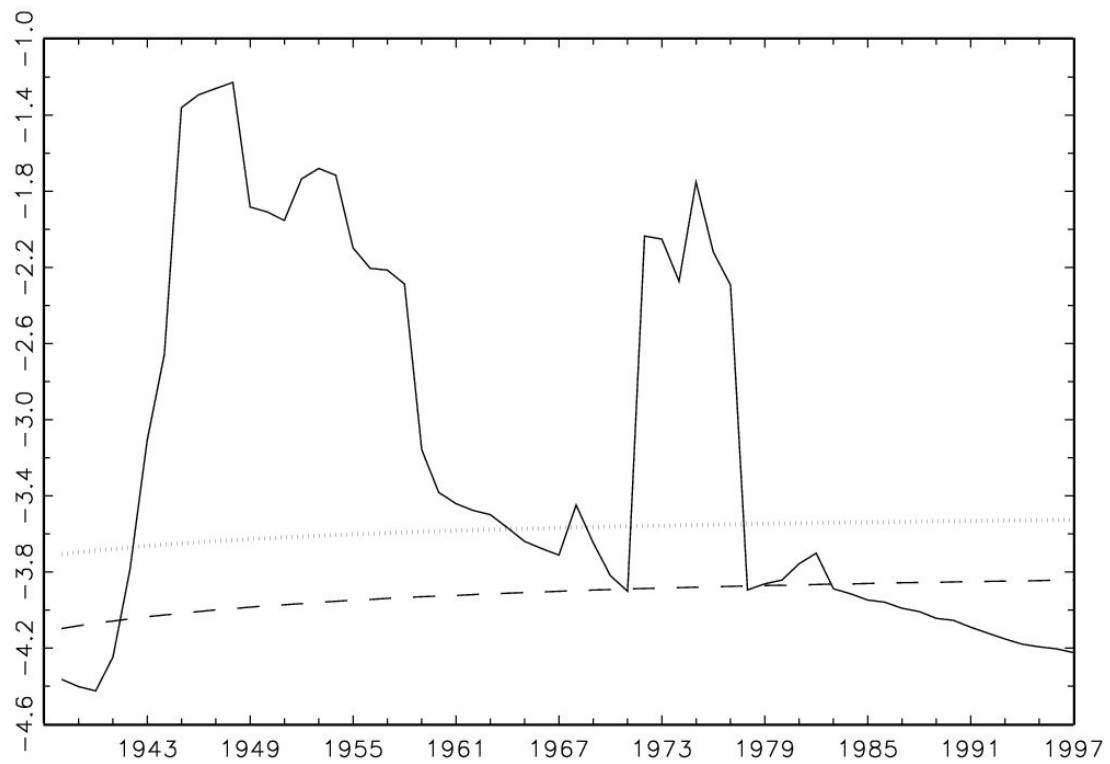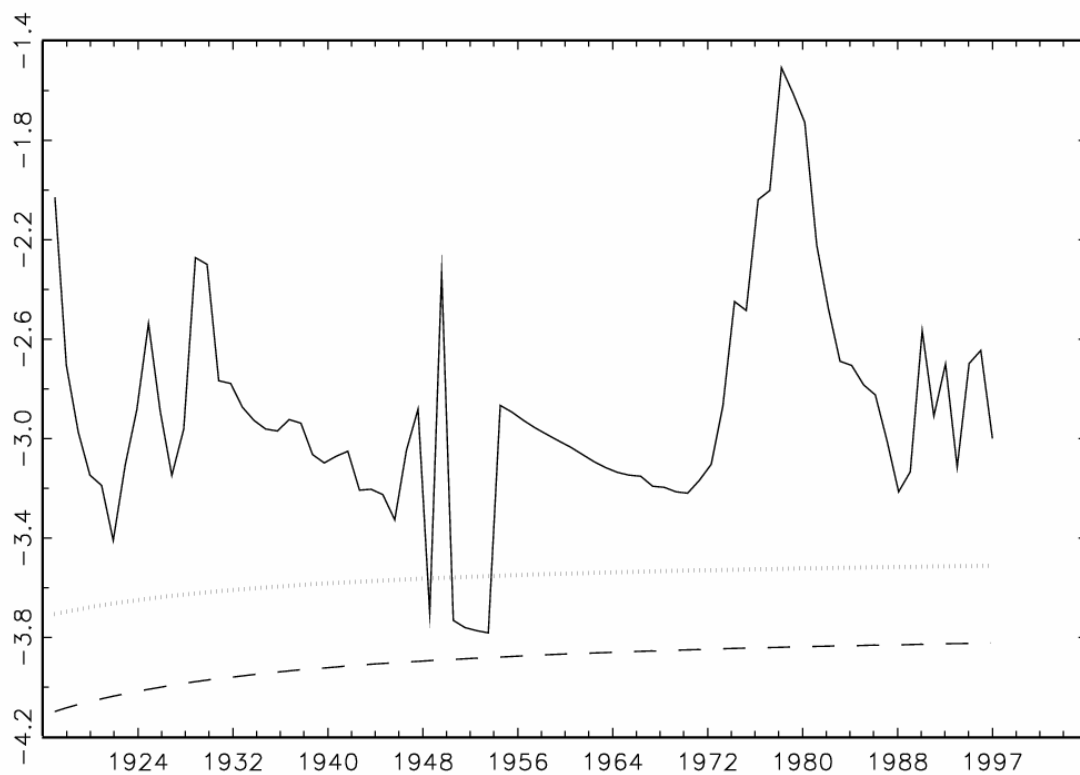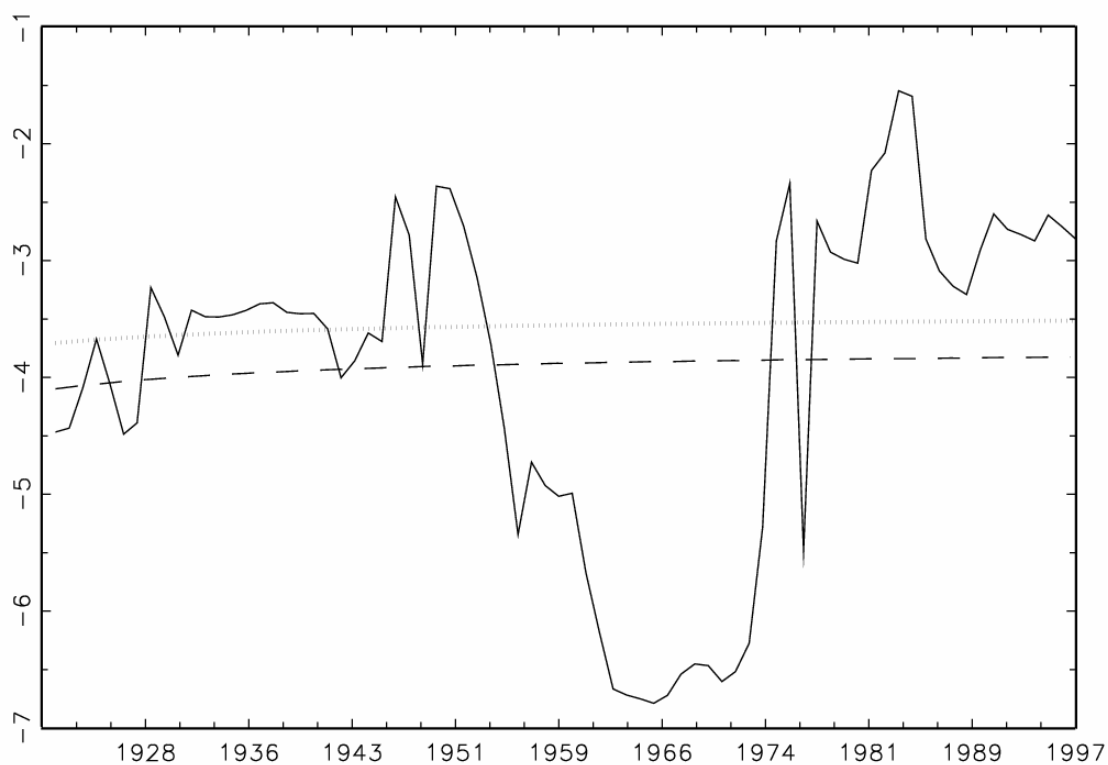

Figure 4: CPI-based Chilean t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression

Figure 5: CPI-based Danish t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression

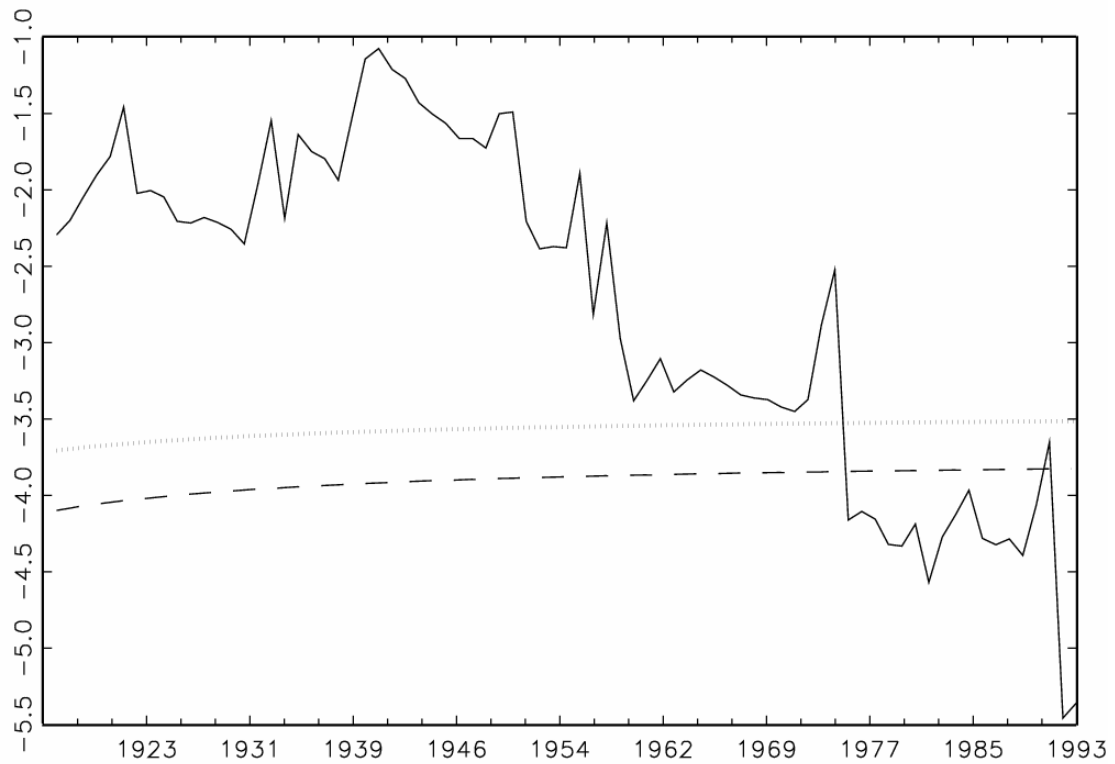**Figure 6: CPI-based Danish t-stat series using a moving window of size $T^* = 30$**

Figure 7: GDP deflator-based Argentinean t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression
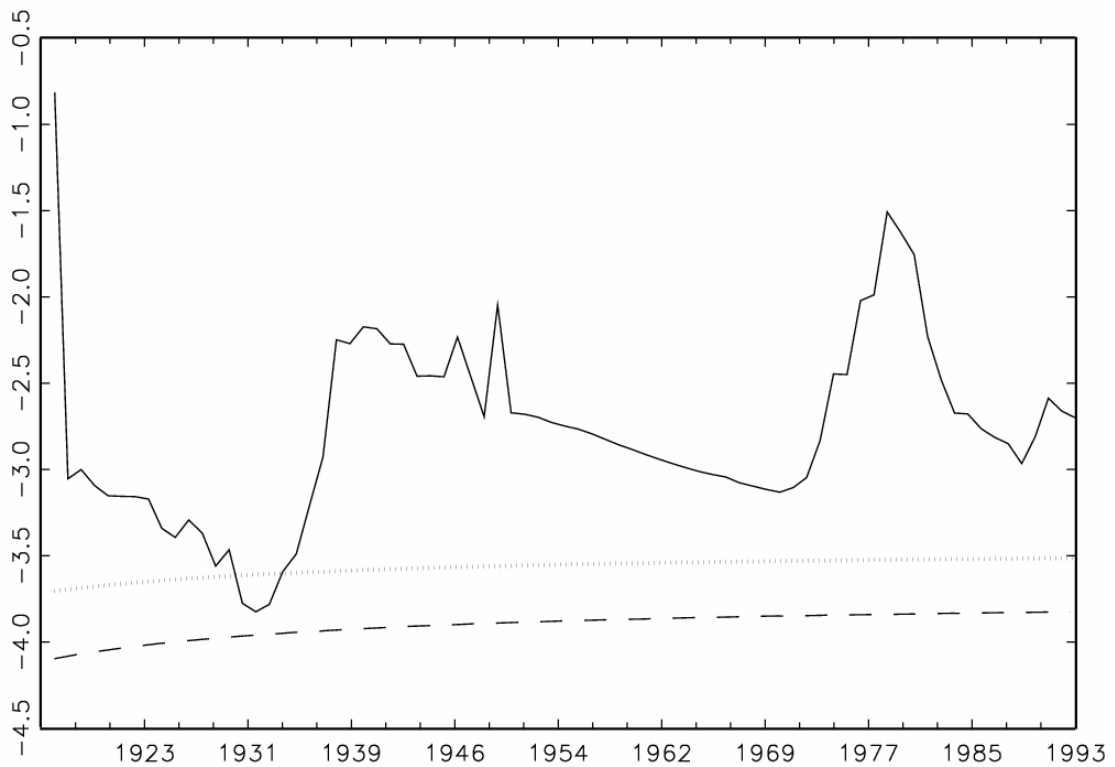

Figure 8: GDP deflator-based Danish t-stat series using a data-dependent rule (Ng and Perron, 2001) for the choice of lags in the ADF regression
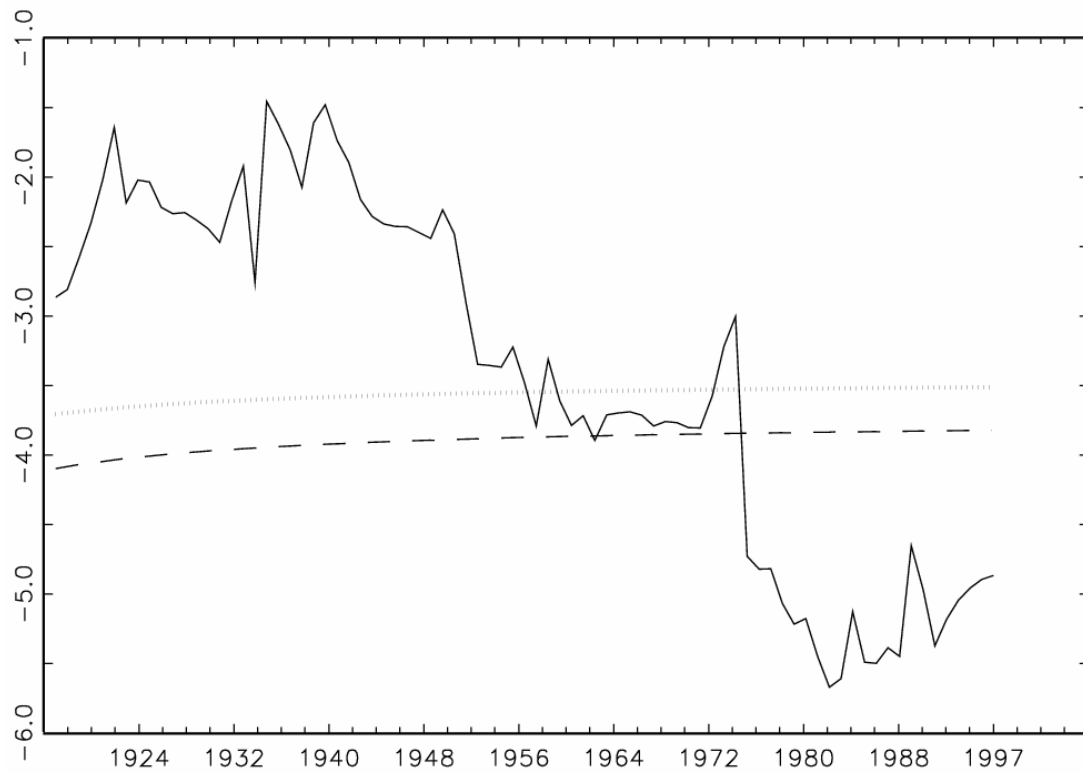
Figure 9: CPI-based Argentine t-stat series using 1 lag in the ADF regression
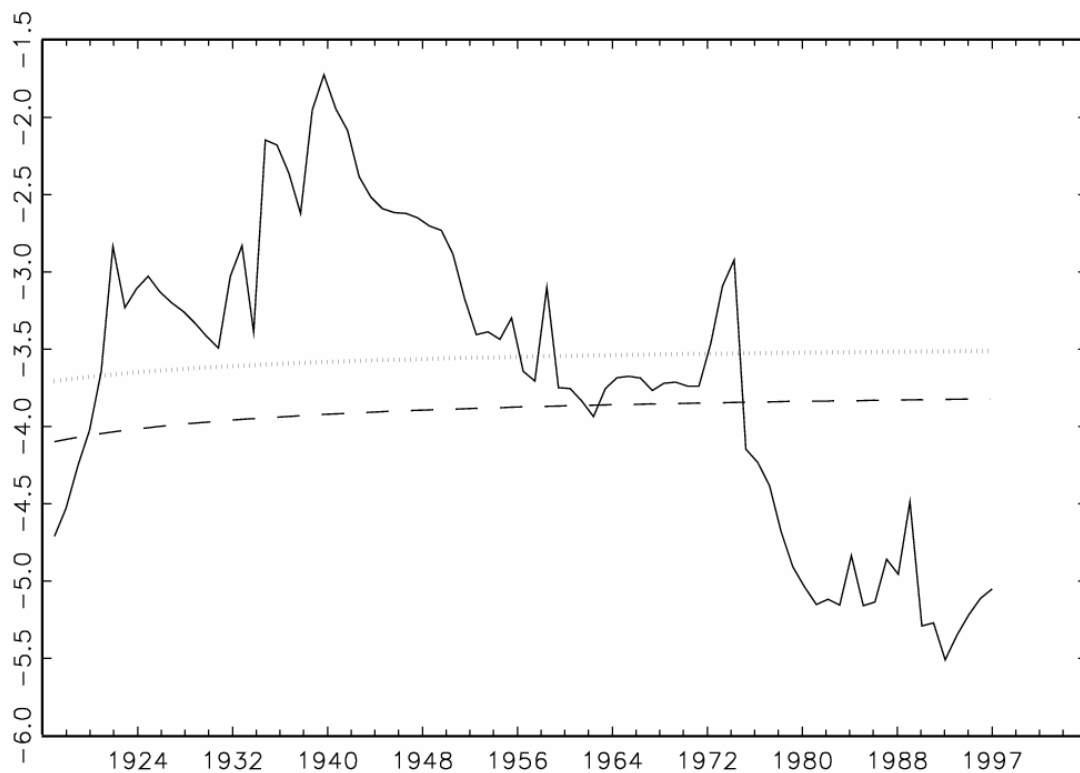

Figure 10: CPI-based Argentine t-stat series using 2 lags in the ADF regression
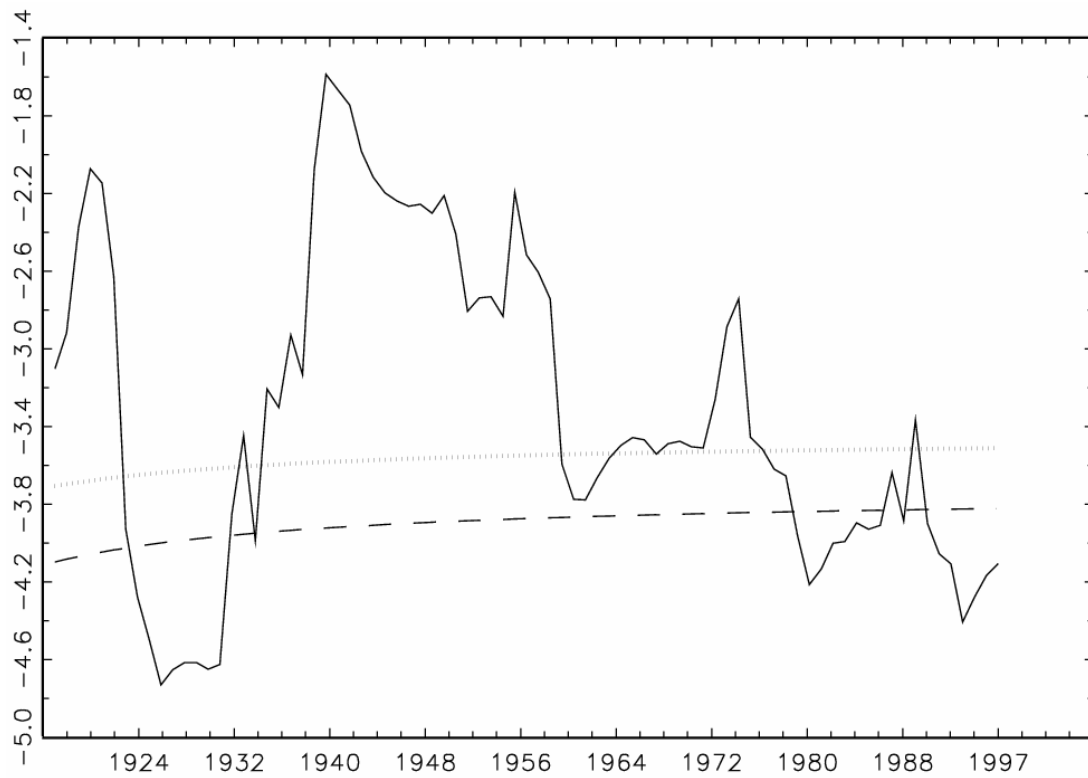
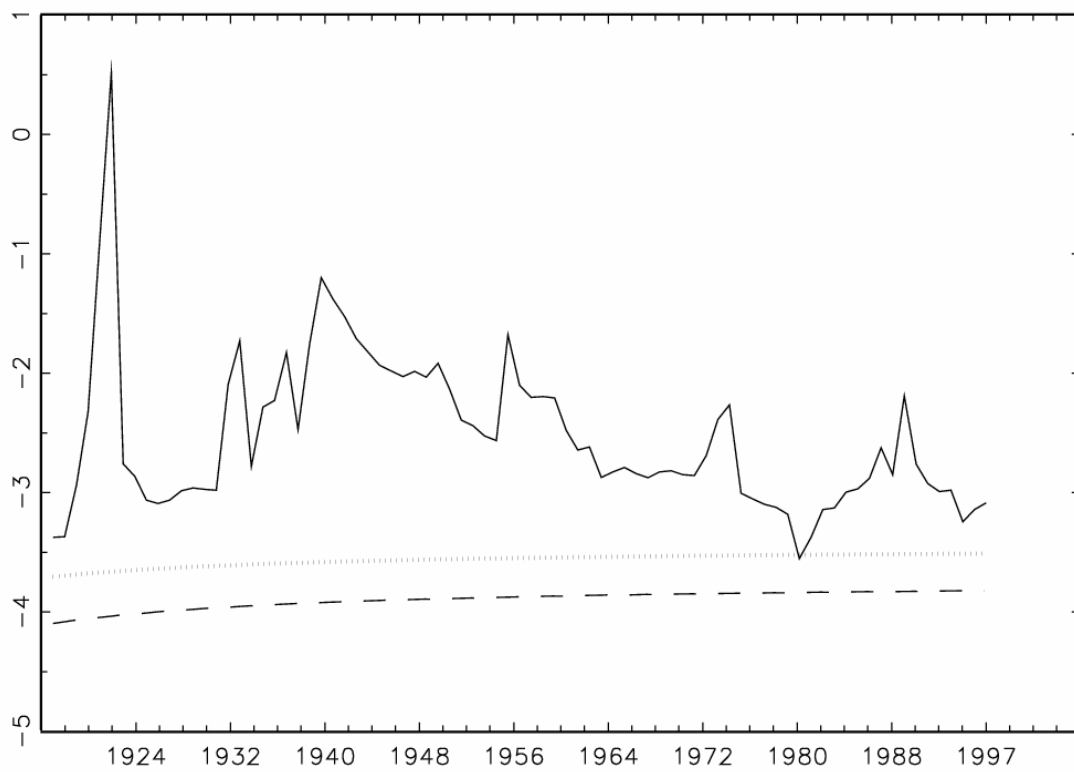**Figure 11: CPI-based Argentine t-stat series using 3 lags in the ADF regression**

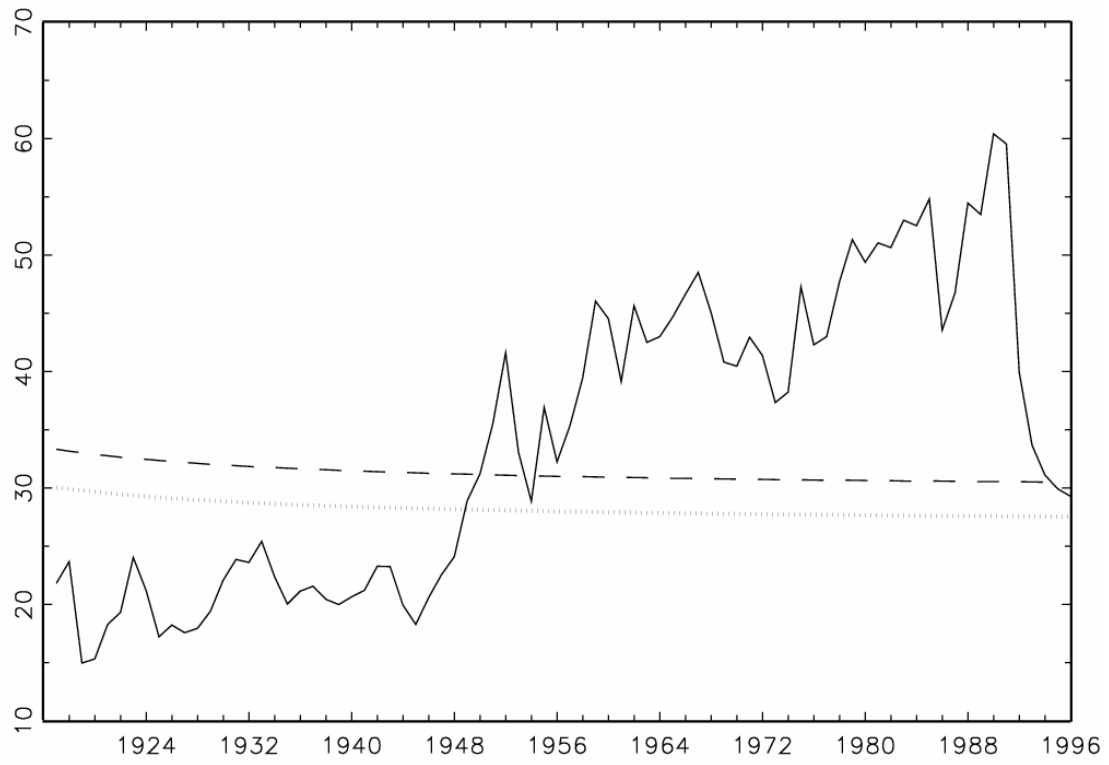**Figure 12: CPI based Argentine t-stat series using 4 lags in the ADF regression**

16

Figure 13: CPI-based Argentine λ -trace stat series

Table 2: Minimum and maximum t-test statistics, acceptance and rejection percentages and number of available observations for each country, using CPI price series

country	Min	Max	Accept	Reject	Obs
ARGENTINA	-5.635	-1.259	0.728	0.272	80
AUSTRALIA	-4.788	-0.536	0.975	0.025	80
BELGIUM	-3.734	0.22	1	0	80
BRAZIL	-2.947	-0.395	1	0	60
CANADA	-4.253	-0.437	0.988	0.012	80
CHILE	-4.427	-1.229	0.65	0.35	59
DENMARK	-3.782	-1.509	1	0	80
FINLAND	-4.797	-0.471	0.481	0.519	80
FRANCE	-5.311	0.204	0.951	0.049	80
GERMANY	-3.867	-1.104	1	0	80
ITALY	-3.666	-1.336	1	0	80
JAPAN	-6.253	-4.109	0	1	24
MEXICO	-5.481	0.289	0.383	0.617	80
NETHERLANDS	-4.092	0.774	0.988	0.012	80
NEW ZEALAND	-5.372	-2.259	0.56	0.44	24
NORWAY	-4.289	-0.496	0.988	0.012	80
PORTUGAL	-5.923	-1.71	0.852	0.148	80
SPAIN	-3.242	-0.018	1	0	80
SWEDEN	-4.219	-1.773	0.852	0.148	80
SWITZERLAND	-3.234	-0.279	1	0	80
UK	-6.642	-1.551	0.802	0.198	80

Table 3: Minimum and maximum λ trace test statistics, acceptance and rejection percentages and number of available observations for each country, using CPI price series

country	Min	Max	Accept	Reject	Obs
ARGENTINA	14.961	60.399	0.45	0.55	80
AUSTRALIA	12.710	54.296	0.9	0.1	80
BELGIUM	28.806	81.093	0.025	0.975	80
BRAZIL	16.048	35.891	0.883	0.117	60
CANADA	22.723	64.069	0.375	0.625	80
CHILE	18.919	40.487	0.322	0.678	59
DENMARK	28.834	75.214	0.025	0.975	80
FINLAND	49.730	75.629	0	1	80
FRANCE	16.919	66.961	0.688	0.313	80
GERMANY	19.392	77.090	0.025	0.975	80
ITALY	27.789	85.518	0.3	0.7	80
JAPAN	34.481	70.718	0	1	24
MEXICO	45.355	96.397	0	1	80
NETHERLANDS	16.44	89.232	0.775	0.225	80
NEW ZEALAND	21.102	48.917	0.458	0.542	24
NORWAY	25.484	81.281	0.175	0.825	80
PORTUGAL	12.222	95.157	0.325	0.675	80
SPAIN	16.991	34.764	0.925	0.075	80
SWEDEN	34.655	111.531	0	1	80
SWITZERLAND	16.630	36.658	0.813	0.188	80
UK	28.242	78.313	0.063	0.938	80