WORKING PAPER

2007 / 1

ESTIMATION BIAS AND INFERENCE IN OVERLAPPING AUTOREGRESSIONS: IMPLICATIONS FOR THE TARGET ZONE LITERATURE

Zsolt Darvas

Budapest, 2007.

This paper is forthcoming in the Oxford Bulletin of Economics and Statistics

Estimation Bias and Inference in Overlapping

Autoregressions: Implications for the Target Zone

Literature*

Zsolt Darvas[†]

February 27, 2007

Abstract

Samples with overlapping observations are used for the study of uncovered interest rate parity, the predictability of long run stock returns, and the credibility of exchange rate target zones. This paper quantifies the biases in parameter estimation and size distortions of hypothesis tests of overlapping linear and polynomial autoregressions, which have been used in target zone applications. We show that both estimation bias and size distortions generally depend on the amount of overlap, the sample size, and the autoregressive root of the data generating process. In particular, the estimates are biased in a way that makes it more likely that the predictions of the Bertola-Svensson-model will be supported. Size distortions of various tests also turn out to be substantial even when using a heteroskedasticity and autocorrelation consistent covariance matrix.

JEL Classification: C22, F31

Keywords: drift-adjustment method, exchange rate target zone, HAC covariance, overlapping observations, polynomial autoregression, size distortions, small sample bias

*This paper is forthcoming in the *Oxford Bulletin of Economics and Statistics*. I thank for comments and suggestions to Christopher Bowdler (editor), Casper G. de Vries, László Halpern, László Hunyadi, Gábor Kőrösi, Helmut Lütkepohl, Judit Neményi, Peter C.B. Phillips, András Simon, András Simonovits, János Vincze, two anonymous referees, and seminar participants at the Winter Symposium of the Econometric Society, the Spring Meeting of Young Economists, and the Econometric Institute of the Erasmus University, Rotterdam. The paper was finalized during my visit to the Tinbergen Institute, Rotterdam. I thank the Tinbergen Institute for its hospitality. I am alone responsible for any errors and for the views expressed in the paper.

[†]Corvinus University of Budapest, Department of Mathematical Economics and Economic Analysis, Fővám tér 8, Budapest, Hungary 1093, E-mail: zsolt.darvas@uni-corvinus.hu, URL: <http://www.uni-corvinus.hu/darvas>

1 Introduction

A time series sample of a regression is said to be overlapping, when the dating of the regressand leads the dating of the regressor by more than one period. There are at least three main areas in empirical macroeconomics and finance where overlapping samples are used: the analysis of uncovered interest rate parity (UIP), the predictability of long run equity returns, and the credibility of exchange rate target zones. In the first two of these three applications, multi-period ahead changes in asset prices are regressed on another variable, that is, the future change in the exchange rate is regressed on the current interest rate differential, and the future change in equity prices is regressed on the dividend yield, respectively. Target zone applications, on the other hand, are based on overlapping autoregressions.

The consequences of overlapping observations for hypothesis tests for the first two applications have already been studied, see, for example, Hansen and Hodrick (1980), Richardson and Stock (1989), Richardson and Smith (1991), Hodrick (1992), Nelson and Kim (1993), Phillips *et. al.* (1996), Smith and Yadav (1996), Kirby (1997), Britten-Jones and Neuberger (2004), and Ang and Bekaert (2005). These papers either suggested methods to estimate the covariance matrix when the sample is overlapping or studied the small sample properties of the estimation and generally reported size distortions of hypothesis tests.¹ Regardless of the covariance matrix used, the consensus of the literature is the empirical failure of UIP. The predictability of long run equity excess returns (over the risk free interest rate) had long been treated as a stylized fact (see, e.g. Fama and French, 1988), but more recent papers showed that this results was due to small sample properties of standard hypothesis tests, and the excess return predictability by the dividend yield is not significant (see Ang and Bekaert, 2005, and references therein).

¹Estimation is usually done by least squares, with the exception of Phillips *et. al.* (1996), who uses the fully modified least absolute deviations (FM-LAD) regression procedure.

To our knowledge, however, the consequence of overlapping observations for estimation and inference in target zone applications, and more generally, in overlapping autoregressions has not yet been studied. Although the issue of overlapping observations is referred to in most of the empirical target zone applications with cautionary warnings, the sole remedy for the problem has been the calculation of the significance of the estimates by the heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator of Newey and West (1987).² This paper aims to fill this gap by studying the properties of overlapping autoregressions along three main dimensions: (a) the amount of overlap, (b) the root of the data generating autoregressive process, and (c) the sample size by Monte Carlo simulation.³ In addition to size properties of hypothesis tests, we also lay special emphasis on the biases of parameter estimations.

In target zone applications the key question is whether the exchange rate within the band is a mean reverting process, and hence, can be forecast adequately based on its past values. When this is the case, then the so called 'drift-adjustment' method can be used, which was suggested by Bertola and Svensson (1993). This method is used to study the possible mean reversion of the exchange rate within the band, the empirical fit of theoretical target zone models, and to predict realignments. The key forecasting equation used for the method is either an overlapping linear autoregression,

$$x_{t+k} = \beta_0 + \beta_1 x_t + v_{t+k}, \quad (1)$$

or a certain non-linear specification, which we call a 'polynomial autoregression',

²This procedure was adopted in, for example, Frankel and Phillips (1992), Bertola and Svensson (1993), Caramazza (1993), Lindberg *et. al.* (1993), Svensson (1993), Rose and Svensson (1994), Thomas (1994), Rose and Svensson (1995), Werner (1996), Chen and Giovannini (1997), Knot, Sturm and de Haan (1998), and Knot and Sturm (1999).

³The distributions can be derived analytically by extending the functional central limit theorem, which was first applied to the problem of deriving the asymptotic distribution of parameter estimates when there is a unit root by Phillips (1987). However, as our main goal is numerical characterization, we do not carry out this exercise in this paper.

$$x_{t+k} = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + v_{t+k} \quad (2)$$

where x_t denotes the time series studied, β_i ($i = 0, 1, 2, 3$) denotes the overlapping autoregressive parameters, and u_{t+k} and v_{t+k} denote the forecast errors, which are assumed to be uncorrelated with information available at time t . The survey presented in Garber and Svensson (1995) regarding this method concluded that exchange rates within the band of target zone currencies displayed strong mean reversion, the empirical fit of the exchange rate function coincided with the predictions of the Bertola-Svensson (BS) model, and estimated devaluation expectations were found to predict actual realignments reasonably well.

The main result of our paper regarding overlapping autoregressions can be summarized as follows. In general, the bias in parameter estimates and size distortions of hypothesis tests are larger, if the amount of overlap is larger, the autoregressive root of the data generating process is closer to unity, and the sample size is smaller. The most notable exception to these general features is the estimator of β_2 in the polynomial autoregression, which is unbiased, though hypothesis tests on this parameter are also affected by size distortions.

Regarding target zone applications, our results indicate that the estimates of linear and polynomial autoregressions are biased in a way that they are likely to overstate support for the theoretical prediction of mean reversion of the expected exchange rate, and other predictions of the BS model, both when the exchange rate is moving in a target zone and when it is floating and following a random walk.⁴ Moreover, size distortions of hypothesis tests make these biased estimated seem-

⁴Mundaca (2004) criticized the drift-adjustment method from a different point of view, claiming that the procedure is burdened with the so called 'peso problem'. However, her two-state Markov-switching model, as well as similar models of Mizrach (1995) and Hallwood *et al.* (2000) are also subject to the issue studied in our paper, since these papers also assume that the exchange rate within the band follows an overlapping autoregression. Estimation is made within a larger model with maximum likelihood, however, there are no reasons to believe that this way of estimating is exempt from the consequences of overlapping observations.

ingly significant.⁵ Using critical values from the distributions simulated in this paper, neither EMS currencies (with the exception of the Dutch Guilder), nor Nordic currencies, nor emerging market currencies display significant mean reversion and non-linear behavior according to the overlapping autoregressions estimated in the literature, but are not significantly different from unit root processes.⁶

The remainder of the paper is organized as follows. Section 2 sets up the economic problem at hand and highlights the main parameters that drive the empirical results. Section 3 presents our Monte Carlo study on the distribution of parameter estimates and hypothesis tests of overlapping linear and polynomial autoregressions. Finally, some concluding remarks are made in Section 4.

2 The drift-adjustment method

In response to reported empirical failures of the target zone model of Krugman (1991), which was the first formal target zone model in the literature and was derived on the basis of perfect credibility, Bertola and Svensson (1993) presented a model with a stochastic time-varying devaluation risk and argued that the existence of this risk can in principle explain empirical regularities of target zone exchange rates. They suggested the drift-adjustment method for extracting a measure of devaluation risk from data, which is also used to validate the theoretical model. Before going into the details of the method, it is worth summarizing the criteria used to evaluate the empirical results. These are the following: (1) some summary statistics of the estimated models (for example, the R^2 and joint F -tests on the parameters), (2) whether the estimated model implies that the exchange

⁵In the context of target zones and unit root tests, Iannizzotto and Taylor (1999) study the power of Dickey-Fuller tests by simulating their estimated target zone model. They find very low power. For example, using a 5 percent critical value they could reject the wrong null-hypothesis of unit root only in very few cases, mostly below 10 percent of experiments.

⁶In a recent paper, Lundbergh and Teräsvirta (2006) offer an elegant solution to the unit root issue by adopting a STAR (smooth transition autoregression) specification. In a non-overlapping model, they allow both the mean and the conditional variance to depend on the within band position. One of their results is, for example, that during July 1985 - May 1991 the Swedish Krona was characterized as a random walk with GARCH variance within the band, but close to the edges of the band the exchange rate behaved like a white noise.

rate within the band displays statistically significant mean reversion, (3) whether the estimates fit the curves implied by the theoretical BS-model, and (4) whether the estimates for expected devaluation predict actual devaluation reasonably well or coincide with the prior conception of the authors.

2.1 Sketch of the method and the first two criteria

By definition, the logarithm of the exchange rate (s_t) managed in an exchange rate target zone can be written as the sum of the logarithm of the central parity (c_t) and the logarithm of the deviation from the central party (x_t), which is usually called 'the exchange rate within the band'. Under the assumption of UIP, the interest rate differential at time t equals the expectations of currency depreciation conditional on information available at time t , which, by definition, equals its two components:

$$r_t - r_t^* = \frac{E_t [\Delta_k s_{t+k}]}{\tau} \equiv \frac{E_t [\Delta_k c_{t+k}]}{\tau} + \frac{E_t [\Delta_k x_{t+k}]}{\tau} \quad (3)$$

where $E_t[\cdot]$ denotes the expectations operator conditional on information available at time t , $\Delta_k s_{t+k} = s_{t+k} - s_t$ is the k -period ahead change of the exchange rate, $k \in N^+$ is the horizon of forward-lookingness, r_t and r_t^* are annualized domestic and foreign interest rates with τ years of maturity, $\tau = \frac{k}{N}$, where N is the number of observations per year.⁷ Hence, by rearranging (3), the expected rate of devaluation can be measured as the difference between the interest differential and the expected rate of depreciation within the band⁸:

⁷For example, in the case of 3-month horizon and 260 working days a year, $k = 65$, $N = 260$, and $\tau = 0.25$. Note that the use of the UIP condition does not necessarily imply overlapping observations for the empirical analysis, namely, when the sample is sequenced in k periods (i.e. using 4 observations per year in case of the 3-month horizon). However, as k -period sequencing substantially reduces the number of observations, empirical studies used to employ data at the available (e.g. daily) frequency and adopt overlapping observations.

⁸Equation (4) shows the essence of the method, as its derivation is slightly complicated by the possible jump of the within band exchange rate upon realignment, see, e.g. Svensson (1993). In practice, this issue is handled by excluding k observations before each case of devaluation (i.e. excluding 3-month data before each devaluation).

$$g_t \equiv \frac{E_t [\Delta_k c_{t+k}]}{\tau} = r_t - r_t^* - \frac{E_t [\Delta_k x_{t+k}]}{\tau} \quad (4)$$

where g_t denotes the expected rate of devaluation.

Bertola and Svensson (1993) suggest that it is sufficient to find an estimate for the expected depreciation within the band, in order to give an estimate of the expected rate of devaluation. The estimate is given by overlapping linear or polynomial autoregressions in the forms of equation (1) or (2). The models are estimated either in levels, or in a k -period differenced form, where both sides are further divided by τ in order to have the annualized measure of the within band depreciation on the left hand side, for example⁹,

$$\frac{x_{t+k} - x_t}{\tau} = \beta_0/\tau + ((\beta_1 - 1)/\tau) x_t + (1/\tau) u_{t+k}. \quad (5)$$

The forecasting equation is occasionally augmented with additional variables, like domestic and foreign interest rates, the interest rate differential, other EMS rates within the band, effective limits, and lags of x_t .

The first two criteria introduced in the beginning of this section are directly related to the estimation of equation (1) or (2). The goodness of fit is usually evaluated by the R^2 and the F -statistics of these overlapping autoregressions, while mean reversion is evaluated by a test of $\beta_1 < 1$, which is done by comparing the t -statistics of the OLS estimate to the Dickey-Fuller table. However, serious doubts could be raised for this procedure on the basis of its application for floating exchange rates: one would find qualitatively the same results as one would find for applications to target zones currencies of the literature that were reported to be favorable. For instance, estimation results of equations (2) and (5) for daily dollar/mark rates for January 2, 1990 - July 31, 2003 are:

⁹Note that equation (5) is equivalent to equation (1).

$x_{t+22} = 0.3 - 0.8x_t + 3.1x_t^2 - 1.7x_t^3 + \hat{v}_{t+k}$, and $(x_{t+65} - x_t)/0.25 = 0.2 - 0.4x_t - 1.5r_t + 1.4r_t^* + \hat{u}_{t+65}$, with all parameters seemingly significant using Newey-West HAC covariances with 21 lags.¹⁰ Hence, overlapping autoregressions would indicate that the logarithm of the dollar/mark rate is stationary, which is in sharp contrast to the consensus of the literature¹¹. Section 3 will demonstrate that the distribution of the F and t tests of overlapping autoregressions are fundamentally different from the χ^2 and Dickey-Fuller distributions, respectively.

2.2 The third criterion: Curves implied by theory

The theoretical BS model defines the so-called 'aggregate fundamental' as the linear combination of the fundamentals of the monetary model of exchange rates ('traditional fundamental') and the expected rate of devaluation. The model specifies the exchange rate function as

$$x_t = h_t + \alpha E_t[\Delta_k x_{t+k}]/\tau \quad (6)$$

where h_t is the aggregate fundamental. Rose and Svensson (1995) estimate the fundamental by

$$\hat{h}_t = \hat{f}_t - c_t + \hat{\alpha}\hat{g}_t \quad (7)$$

where $\hat{f}_t = s_t - \hat{\alpha}(r_t - r_t^*)$ is an estimate of the 'traditional fundamental', $\hat{\alpha}$ is an estimate of the interest rate semi-elasticity of money demand, and $\hat{g}_t$ is the estimate of the expected rate of devaluation by the drift adjustment-method based on equation (4), using an estimate of (1) or

¹⁰The starting year 1990 was selected to exclude the period of the Plaza Agreement and the Louvre Accord when interventions in the main foreign exchange markets were (or were thought to be) more intensive than before or afterwards. For the period since the introduction of the euro in 1999, German mark exchange rates were calculated from the euro rates using the conversion rate, and German interest rates values were substituted with euro interest rates. Results were qualitatively insensitive to the specific sample period chosen. We also studied various other floating nominal exchange rates with qualitatively similar results, which are available from the author upon request.

¹¹See, for example, Frankel and Rose (1995). Using the dataset of this paper we also run unit root and stationarity tests for various floating nominal exchange rates, which all indicated non-stationarity. Results of these tests are available from the author upon request.

(2) to form an estimate of the expected change within the band, $\widehat{E}_t[\Delta_k x_{t+k}]$. Substituting these equations into (7) leads to the estimate: $\widehat{h}_t = x_t - \widehat{\alpha} \widehat{E}_t[\Delta_k x_{t+k}]/\tau$, which is, of course, identical to the method of Rose and Svensson (1995) and can be viewed as a rearrangement of equation (6).

*** Figure 1 ***

Theoretical predictions of the BS model are illustrated in Panels A-B-C of Figure 1. Panel A plots the expected depreciation within the band against the exchange rate within the band. The slope of the locus is negative and its curvature is convex towards the strong edge and concave towards the weak edge. The vertical difference between the curve and the interest rate differential is regarded as the expected rate of devaluation according to equation (4). Panel B plots the expected exchange rate within the band against the exchange rate within the band and the 45° line. Panel C plots the exchange rate within the band against the aggregate fundamental and the 45° line. The theoretical model suggests an exchange rate honeymoon effect in terms of the aggregate fundamental, that is, a unit shock to the aggregate fundamental cause less than a unit shock to the exchange rate within the band. Panel C also plots x_t against the traditional fundamental scaled down by the central parity. The horizontal difference between this value and the new exchange rate function is regarded as α times the expected rate of devaluation according to equation (7).

2.3 Parameters driving the curves implied by theory

There is a direct relationship among the three figures of the BS model. First, it is clear that Figure 1A implies Figure 1B, since Figure 1A and 1B show the change and the level of x_{j+k} , respectively. Second, we showed that the aggregate fundamental could be calculated as $\widehat{h}_t = x_t - \widehat{\alpha} \widehat{E}_t[\Delta_k x_{t+k}]/\tau$. Since α (and its estimate $\widehat{\alpha}$) is positive, provided that the empirical counterpart of Figure 1A has the shape displayed, plotting $(x_t - \widehat{\alpha} \widehat{E}_t[\Delta_k x_{t+k}]/\tau, x_t)$ reveals the exchange rate honeymoon effect in terms of the aggregate fundamental. Thus, Figure 1A implies both of the other two figures implied by the target zone theory. Let us highlight the main parameters driving the plot of Figure

1A, that is, the curvature of $E_t[\Delta_k x_{t+k}]$ against x_t .

The estimated model corresponding to the theory behind the curve in Figure 1A is a third order polynomial, where x ranges between the lowest and highest possible values, for example, between -2.25 percent and +2.25 percent in the case of target zone currencies and between 0.30 and 0.86 in the case of the logarithm of the dollar/mark rate. Every third order polynomial goes to plus and minus infinity at the different ends of the axis, and the signs are determined by the parameter of the cubic term. It is also clear that there might or might not be a hump somewhere in the middle of the domain. There is a hump only if the first derivative changes sign. This in turn depends on the roots of the polynomial, whether there are three distinct real roots or not.

Let $y_t = x_{t+k} - x_t$ and subtract x_t from each side of equation (2). For the sake of simplicity, let us drop the time index. The resulting polynomial is the function plotted on Figure 1A,

$$y = \beta_0 + (\beta_1 - 1)x + \beta_2 x^2 + \beta_3 x^3. \quad (8)$$

Regardless of β_1 , the second derivative changes sign at $x = -\beta_2/(3\beta_3)$ (provided, of course, that $\beta_3 \neq 0$). This means that the curvature changes from concavity to convexity or vice versa along the axis. In the case of a negative β_3 the value of the polynomial (8) approaches infinity and minus infinity in the left and right direction, respectively. A continuously negative first derivative ensures the absence of a hump in the middle of the domain. Thus, provided that $\beta_3 < 0$ and

$$(\beta_1 - 1) + 2\beta_2 x + 3\beta_3 x^2 < 0, \quad (9)$$

the slope of the line of Figure 1A will be negative without any hump in the whole domain with curvature implied by the BS theory. Condition (9) would be more likely fulfilled if β_1 was less than one, β_2 was close to zero, and β_3 was negative. However, these are the general properties of

estimation results for floating exchange rates, and in Section 3 we will demonstrate that these are exactly the results that one obtains from Monte Carlo simulation for random walks and near unit root processes when studying the statistical properties of the estimation. As an example, Panels D to I of Figure 1 show the estimated curves for the log-levels of two floating exchange rates. The remarkable feature of these panels is the similarity to the curves implied by the BS theory and to the empirical estimates for real world target zone data (see, for example, Figures 10 and 11 of Bertola and Svensson, 1993, and Figures 1 and 2 of Rose and Svensson, 1995).¹²

3 Bias and inference in overlapping autoregressions

It is well known, at least since the work of Dickey and Fuller (1979), that in a simple first order autoregression the OLS estimate of the autoregressive parameter is downward biased in small samples both for the stationary and the unit root cases. The limiting distribution of the deviation of the estimated coefficient from the true value multiplied by the convergence rate is symmetric in the stationary case [that is, $\sqrt{T}(\hat{\rho} - \rho)$] but asymmetric in the unit root case $[T(\hat{\rho} - 1)]$, with negative values being much more likely, where T denotes the sample size, ρ the population value of the autoregressive parameter, and $\hat{\rho}$ is the least squares estimator. In this section we describe various small and large sample distributions of parameters and hypothesis tests of overlapping linear and polynomial autoregressions which, to our knowledge, have not been studied before. Our main goal is to study the inference and, particularly, to assess the biases and size distortions arising from the estimation using overlapping observations in small samples. We will use our findings to evaluate the significance of results achieved for the application to target zone currencies of the literature and to the floating exchange rates presented in this paper.

¹²The fourth criterion, realignment prediction, is also related to the biases in the estimated parameters of the depreciation equation. Specifically, when condition (9) is satisfied by the estimated parameters, then a weaker exchange rate within the band implies higher realignment probability. As a weaker exchange rate within the band usually precede cases of devaluation which are followed by a stronger within the band exchange rate, it is not surprising that inferred devaluation expectations have some correlation with actual devaluations.

It can be proven analytically, by extending the functional central limit theorem, that, when the true process is a random walk, then the OLS parameter estimator of both the linear and the polynomial overlapping autoregressions is consistent. In the case of the polynomial autoregression the convergence rate of $(\hat{\beta}_1 - 1)$ is T , similar to the simple linear autoregressive model when there is a unit root. The convergence rate of $(\hat{\beta}_2 - 0)$ is $T^{3/2}$, and the convergence rate of $(\hat{\beta}_3 - 0)$ is T^2 . Moreover, only the distribution of $T(\hat{\beta}_1 - 1)$ is independent of the residual variance, but both $T^{3/2}\hat{\beta}_2$ and $T^2\hat{\beta}_3$ depend on that.¹³ Since small sample biases also depend on the residual variance, in our simulation study we set it equal to the variance of daily logarithmic changes of the dollar/mark rate, which is 0.006821².

We study the properties of the estimators and hypothesis tests along three dimensions: (a) the amount of overlap, (b) the root of the data generating autoregressive process, and (c) the sample size.¹⁴ As a benchmark, we set k , the amount of overlap to 22 (representing the working days in a month); T , the sample size to 3000 (the typical sample size used for EMS currencies in the literature and for floating exchange rates in our paper); and λ , the inverted root of the data generating autoregressive process to 1 (implying random walk). In the next three subsections we study the properties of the OLS estimator and various hypothesis tests along these three main dimensions.

Our main emphasis will be the investigation of the effects of the amount of overlap for stationary processes as well as for unit root processes. The issue of unit root is crucial from the point of view of hypothesis testing: whether estimated models differ from the random walk, or to phrase it

¹³The proofs are available from the author upon request.

¹⁴We also briefly studied the effects of shifts in the intercept, because it was common to adjust for different exchange rate regimes this way to account for change in the mean of the exchange rate process. Common sense would suggest that once we allow the intercept to change, we need a lower estimate of the autoregressive parameter in order to be able to reject the null hypothesis of a unit root in favor of stationarity. This was confirmed by our simulation study. For example, in the simple Dickey-Fuller test equation (which is not overlapping), the 5% critical value, -2.86, is switched to -4.32 when there are 5 'regimes' (4 changes in the intercept) and to -5.42 when there are 10 'regimes' (9 changes in the intercept), where we used the sample $T = 3000$ for $n = 10000$ trials and assumed that the regimes are of equal length.

differently, whether the mean reversion and non-linear effects are significant or not.

3.1 The effects of the amount of overlap

We studied the effects of overlapping observations for estimation and inference by estimating the linear and polynomial autoregressions for random walks with $k = 1, \dots, 50$, using a sample of $T = 3000$, which corresponds to the typical sample size used in the literature. Figure 2 shows kernel density estimates of the estimators for five values of k . Panel A shows the density of the autoregressive parameter of the linear autoregression (which is hereafter denoted by $\beta_1^{(AR)}$), while the other three panels show the densities of the parameters of the polynomial autoregression. It is clearly evident that the variance of the estimates increases with the amount of overlap. Furthermore, the distribution of $\hat{\beta}_2$ is symmetric, while the distributions of the other three parameters are highly asymmetric with less than the true values being much more likely.

*** Figure 2 ***

*** Figure 3 ***

Panels A and B of Figure 3 show some numerical characteristics of the densities. Panel A of Figure 3 shows the percentage of estimates with $\hat{\beta}_1$ less than one (both for the linear and polynomial autoregressions) and the percentage of estimates with negative $\hat{\beta}_2$ and $\hat{\beta}_3$ for the polynomial autoregression. Panel A also shows the mean R^2 of the polynomial autoregression (the R^2 of the linear autoregression is similar). It has been shown for linear non-overlapping autoregressions that the probability of an autoregressive parameter estimate less than one is very high, close to 95%. Our results indicate that overlapping observations have virtually no effect on this probability. Recall that the true values of the parameters used for the data generating process are $\beta_1 = 1$, $\beta_2 = 0$, and $\beta_3 = 0$. The amount of overlap also does not affect much the probabilities of getting less than true values for the parameters of the polynomial autoregression, but it is important to highlight the magnitudes of these probabilities: $P(\hat{\beta}_1 < 1) \approx 0.75$, $P(\hat{\beta}_2 < 0) = 0.5$, and $P(\hat{\beta}_3 < 0) \approx 0.92$.

Note that all floating exchange rate estimations of this paper yielded a negative value for $\hat{\beta}_3$, while $\hat{\beta}_2$ had mixed signs.

Panel B of Figure 3 shows the mean bias in parameter estimation. Clearly, the bias in the estimation of β_2 is zero and not related to the amount of overlap, but the bias is systematically related to the amount of overlap in the cases of $\hat{\beta}_1$ (both for the linear and polynomial autoregressions) and $\hat{\beta}_3$. That is, increasing the amount of overlap increases the magnitude of the bias in the estimation of these parameters. Note that the downward bias in $\hat{\beta}_3$ is much more severe than in $\hat{\beta}_1$ using the dollar/mark rate to calibrate the data generating process.¹⁵

The finding that increasing the amount of overlap does not affect the probability of getting an estimate less than the true value but does increase the estimation bias is the consequence of changes in the shapes of the estimator distributions. Figure 2 showed that increasing the amount of overlap substantially expands the left tails of the distributions, while the right tails are only mildly affected.

Panel C of Figure 3 shows the sizes of t -tests, using 5% nominal sizes, calculated with Newey-West HAC covariances with lags equal to $k - 1$, since the order of autocorrelation is $k - 1$ in the case of overlapping observations (see, for example, Hansen and Hodrick, 1980). For $\hat{\beta}_1$ we use the one sided Dickey-Fuller critical value of -2.86, while for $\hat{\beta}_2$ and $\hat{\beta}_3$ we use the two-sided critical value of the t -distribution. For the non-overlapping case (i.e. $k = 1$), our results of course confirm the Dickey-Fuller distribution in the case of the linear autoregression, since the empirical size of the null hypothesis of $\beta_1^{(AR)} = 1$ equals the nominal size of 5%. However, even a single period overlap ($k = 2$) causes a jump in the empirical size to 13%, which increases further with the amount of overlap.

We also see a jump in the empirical sizes when moving from the non-overlapping to the overlap-

¹⁵We have already highlighted that the distributions of $\hat{\beta}_2$ and $\hat{\beta}_3$ depend on the residual variance, while that of $\hat{\beta}_1$ does not.

ping sample in the cases of the parameters of the non-linear terms in the polynomial autoregression. Interestingly, size distortions do not increase further for $k > 5$, but empirical sizes stay fairly constant at around 29% and 38% in the cases of $\hat{\beta}_2$ and $\hat{\beta}_3$, respectively. Among the four parameter estimates $\hat{\beta}_1$ of the polynomial autoregression has the best size properties, although its empirical size also increases with the amount of overlap.

Finally, Panel D of Figure 3 shows sizes of joint hypothesis tests on the parameters of the polynomial autoregression. W1 denotes the null hypothesis of the true process (i.e. $H_0: \beta_0 = \beta_1 - 1 = \beta_2 = \beta_3 = 0$), while W2 denotes the null hypothesis of linearity (i.e. $H_0: \beta_2 = \beta_3 = 0$). Sizes were evaluated using the 5% critical value of the χ^2 distribution, which was also used for testing non-linearity in target zone applications in the literature with Newey-West HAC covariances to correct for autocorrelation.¹⁶ Our results indicate that empirical sizes are substantially higher than nominal sizes, which are also increasing in k . We can also observe a jump in the empirical sizes when moving from the non-overlapping to the overlapping sample.

*** Figure 4 ***

Figure 4 shows the same statistics for a stationary data generating process, $x_t = 0.9x_{t-1} + \varepsilon_t$. Results are qualitatively similar to the unit root case although magnitudes are different. In the linear autoregression, the probability of an estimate less than the true population value is fairly constant at 55% irrespective of the amount overlap, but the magnitude of the bias is increasing in k . Results for $\hat{\beta}_2$ are identical to the unit root case, since $P(\hat{\beta}_2 < 0) = 0.5$ and the estimate is unbiased. Results for $\hat{\beta}_3$ are similar again to the unit root case, since $P(\hat{\beta}_3 < 0) > 0.5$ and the estimate is downward biased, although for large amounts of overlap the bias is diminishing. Interestingly, the effects on the estimate of $\hat{\beta}_1$ changes sign as a function of the amount of overlap.

¹⁶It is well known that in the case of stationary regressors, including stationary autoregression when the regressor(s) is(are) not independent from the population residual variable, the Wald type test has an exact or asymptotically χ^2 distribution with degrees of freedom equal to the number of restrictions. See, for example, Hamilton (1994) p. 461-472.

For $k < 12$ the estimate is upward biased while for $k > 12$ the sign is reversed. For large k , results are qualitatively similar to the unit root case, since the probability does not increase further but the bias is increasing. Hence, the only qualitative difference to the unit root case is the bias in $\hat{\beta}_1$ for small amounts of overlap. However, recall the condition formulated in equation (9), which jointly depends on the parameters and note that the biases in $\hat{\beta}_1$ and $\hat{\beta}_3$ are of different orders of magnitude, as is shown on Panel B of Figure 4. Also note that target zone applications usually adopted an overlap of one month (corresponding to $k = 22$, which is the number of working days in a month) or even a higher amount of overlap. Finally, Panels C and D of Figure 4 indicate that there are no size distortions of hypothesis tests in the non-overlapping case (i.e. $k = 1$), but size distortions are present even in the stationary case when the sample is overlapping.

To sum up, probabilities of getting estimates less than the true value do not depend on the amount of overlap, but biases in parameter estimates (except $\hat{\beta}_2$) and sizes of the hypothesis tests are systematically related to the amount of overlap. $\hat{\beta}_1$ (both in the linear and polynomial autoregressions) and $\hat{\beta}_3$ are downward biased, while $\hat{\beta}_2$ is unbiased.

3.2 The effects of the autoregressive root

We study the effect of the autoregressive root of the data generating process on estimation and inference in more detail by applying the linear and polynomial autoregressions to AR(1) processes with different roots. A zero mean AR(1) process could be written in the form of $(1 - \lambda L)x_t = \varepsilon_t$, where λ denotes the inverted root (which must be inside the unit circle for covariance stationarity), L is the lag operator and ε_t is a white noise process. We set λ to take various values between 0.80 and 1.00. For each process we have studied the properties of overlapping autoregressions with $k = 22$ (corresponding to one month) and $T = 3000$ (corresponding to the typical sample size used in the literature).

*** Figure 5 ***

Panel A of Figure 5 shows the R^2 and probabilities of parameter estimates less than the true population values.¹⁷ The autoregressive parameter of the linear overlapping autoregression indicates a systematic relation to the root of the process: as the root approaches one, the likelihood of a less than true estimate is slightly increasing, but jumps in the unit root case to about 95%, as it was already shown by Figure 3. We can also see a clear pattern for β_3 and β_2 : as the root approaches unity, the estimate for β_3 is more and more likely to be negative (but without a distinct jump at $\lambda = 1$), while the signs of the estimate for β_2 are balanced. However, the sign of $\hat{\beta}_1 - \lambda^{22} < 0$ of the polynomial autoregression is not monotonically related to λ : it is balanced till about $\lambda = 0.92$, then the probability is slightly decreasing till $\lambda = 0.99$ to 42%, but jumps to 75% in the unit root case.

The non-monotonic effects of λ on $\hat{\beta}_1$ in the overlapping polynomial autoregression are also reflected in its bias, which is shown in Panel B. The bias is slightly negative for $\lambda < 0.94$, positive for $0.95 \leq \lambda \leq 0.99$, then jumps to about -0.09 in the unit root case as is shown in the previous figure. The bias in $\hat{\beta}_3$ is also non-monotonic in λ , although it is always negative. Note the scaling of the right axis, which highlights that the bias in $\hat{\beta}_3$ is even more severe for stationary processes having roots between 0.85 and 0.99 than in the unit root case. The bias in $\hat{\beta}_2$ is zero and not related to the root of the process. Finally, the bias in $\hat{\beta}_1^{(AR)}$, the overlapping autoregressive parameter of the linear model, is systematically related to the root of the process, namely, the magnitude of the downward bias is increasing in λ .

Panel C of Figure 5 shows the sizes of individual t -tests on the parameters using Newey-West HAC covariances (with 21 lags) at 5% nominal size. Note that the process is stationary for $\lambda < 1$, hence, we tested the hypothesis of $\beta_1 = \lambda^{22}$ for $\lambda < 1$ with a two-sided t -test, and $\beta_1 = 1$ for

¹⁷The true value of the overlapping autoregressive coefficient equals λ^k , which is also equal the k^{th} -order autocorrelation coefficient. Its magnitude can take values close to zero for $\lambda < 1$. Hence, the overlapping sample is the reason for the almost zero R^2 in the case of stationary processes with low roots.

$\lambda = 1$ with a one-sided Dickey-Fuller test. Panel C indicates that size distortions are also present in a stationary setup. One would reject the true null hypotheses of $\beta_1 = \lambda^{22}$, $\beta_2 = 0$, and $\beta_3 = 0$ too often. It is further noteworthy that empirical sizes are increasing in λ . In the unit root case, empirical sizes are very large.

Finally, Panel D shows the empirical sizes of joint tests on the parameters of the polynomial autoregressions at 5% nominal size, which are also increasing in λ , confirming distortions in a stationary setup. For example, when $\lambda = 0.98$, the chance of rejecting the true null hypothesis of the process (i.e. $H_0: \beta_0 = \beta_1 - 0.98^{22} = \beta_2 = \beta_3 = 0$) is about 29%, and we would reject the true null hypothesis of linearity ($H_0: \beta_2 = \beta_3 = 0$) in 20% of cases, using 5% nominal sizes and Newey-West HAC covariances (with 21 lags) in both cases.

Hence, results in this subsection also support our earlier findings that the OLS estimator of $\hat{\beta}_2$ is symmetric and unbiased, while that of $\hat{\beta}_1$ and $\hat{\beta}_3$ are skewed to the left, and there are considerable size distortions for all hypothesis tests considered.

3.3 The effects of the sample size

Table 1 shows the effects of the sample size on estimation and inference of the overlapping linear and polynomial autoregressions by estimating them for random walks for various sample sizes between 100 and 100000, with $k = 22$.

*** Table 1 ***

We have seen in the previous subsections that $\hat{\beta}_2$ was unbiased, while $\hat{\beta}_1$ (both for the linear and polynomial autoregressions) and $\hat{\beta}_3$ were downward biased. These properties hold for a range of sample sizes. Again, the bias of $\hat{\beta}_3$ is more severe than that of $\hat{\beta}_1$, although this particular result depends on the residual variance, which was calibrated here to the dollar/mark exchange rate. We have also already seen on Figure 3 that $P(\hat{\beta}_1^{(AR)} < 1) \approx 0.95$, $P(\hat{\beta}_1 < 1) \approx 0.75$, $P(\hat{\beta}_2 < 0) = 0.5$, and $P(\hat{\beta}_3 < 0) \approx 0.92$, which are also generally robust to sample size, with perhaps the exception

of $\hat{\beta}_3$ whose probability of being less than the true value is somewhat increasing.

The sample size has three distinct effects. First, biases go to zero as the sample size increases, and the variance of biases also diminishes, which can be seen in the 6th – 9th columns of the upper block of the table containing mean absolute biases. Values in the upper block of the table also indicate that $\hat{\beta}_3$ has the highest convergence rate and $\hat{\beta}_1$ has the slowest. Second, the R^2 increases with the sample size which reflects the fact that biases diminish with the sample size.

The third distinct effect of the sample size is related to the sizes of hypotheses tests. Although there are considerable size distortions at all sample sizes, the empirical size improves for the t -test of $\hat{\beta}_1$ (both for the linear and polynomial autoregressions) and for the joint hypothesis tests (using 5 percent nominal size and Newey-West HAC covariances with 21 lags in all cases). For instance, at the shortest ($T = 100$) and longest ($T = 100000$) sample sizes shown in the table, the empirical sizes of the Dickey-Fuller t -tests in the linear overlapping ($k = 22$) autoregressions are 62% and 16%, respectively (the nominal size is 5%). The empirical sizes of the two joint hypothesis tests are 99% and 55%, and 45% and 30%, respectively, at the shortest and longest sample sizes shown.

The size distortion of t -test of $\hat{\beta}_1$ in the polynomial autoregression seems to diminish, since the empirical size decreases to 5.3% at 5% nominal size. However, this is a mere consequence of the selected amount of overlap and not an indication that Dickey-Fuller tests are valid asymptotically. For instance, at $T = 100000$, the true sizes using 1% and 10% nominal sizes are 1.8% and 8.8%, respectively. Note also that Figure 3 indicated that the true size of the null hypothesis of $\hat{\beta}_1 = 1$ increases with the amount of overlap for a given sample size.

Size distortions of t -tests of $\hat{\beta}_2$ and $\hat{\beta}_3$ are also considerable, reaching 29% and 38%, respectively, for longer samples. However, after an initial increase when moving from very small samples to $T = 1000$, distortions do not increase further with the sample size. Note that this result bears some similarity to results in the previous subsection studying the effect of the amount of overlap,

where we found that, after an initial increase in size at low amounts of overlap, size distortions do not increase further for $k > 5$.

To sum up, probabilities of getting estimates less than the true values do not depend much on the sample size, but the magnitudes of the bias in parameter estimates (except $\hat{\beta}_2$) depend, since estimates are consistent. Sizes of the hypothesis tests also depend on the sample size. Thus, these findings strengthen the generality of the results shown in the previous subsections: $\hat{\beta}_1$ (both in the linear and polynomial autoregressions) and $\hat{\beta}_3$ are downward biased, while $\hat{\beta}_2$ is unbiased, and hypothesis tests have bad size properties even when using the Newey-West HAC covariances.

3.4 Summary and implications for the target zone evidence

Table 2 presents a summary of our main findings along three aspects: the shape of the distribution function, bias, and size distortions of hypothesis tests. Our main results are that the distributions of the estimator of $\beta_1^{(AR)}$, β_1 , and β_3 are skewed to the left and the estimates are downward biased while the distribution of the estimator of β_2 is symmetric and its estimate is unbiased. The empirical size of hypothesis tests on all four parameters, and also of joint hypothesis tests on the parameters are all above the nominal sizes. In general, the bias in parameter estimates and size distortions of hypothesis tests are larger, if the amount of overlap is larger, the autoregressive root of the data generating process is closer to unity, and the sample size is smaller.

*** Table 2 ***

Turning to the target zone applications of the literature and floating exchange rate illustrations of this paper, recall that the validation of the theoretical predictions of the Bertola-Svensson model can be summarized by condition (9). This condition is more likely fulfilled if β_1 was less than one, β_2 was close to zero, and β_3 was negative. However, these are exactly the results what we obtained by Monte Carlo simulation for random walks and near unit root processes. We conclude that both biases of parameter estimates and size distortions of hypothesis tests of the overlapping polynomial

autoregression are such that processes having large autoregressive roots seemingly significantly fit all theoretical predictions of the BS model.

Comparing the 5% critical values of the t -tests and the Wald-tests that are simulated in this paper, neither the floating exchange illustrations of this paper, nor target zone results of the literature¹⁸ for the EMS (with the exception of the Dutch Guilder), Nordic currencies, and emerging market currencies display significant mean reversion. Moreover, neither our floating exchange rate illustrations, nor the FF/DM application of Rose and Svensson (1995) show significant non-linear effects.¹⁹

4 Concluding remarks

This paper studied the effects of overlapping observations on estimation and inference for autoregressions. Samples with overlapping observations are frequently used, for example, for the study of UIP, the predictability of long run stock returns, and the credibility of exchange rate target zones. We took an example from the target zone literature, the so-called drift-adjustment method, which is based on linear and polynomial autoregressions estimated on overlapping observations.

We characterized the properties of parameter estimates and hypothesis tests of the overlapping linear and polynomial autoregressions by Monte Carlo simulation along three main dimensions: (a) the amount of overlap, (b) the root of the data generating autoregressive process, and (c) the sample size. We showed that the bias in parameter estimates and size distortions of hypothesis

¹⁸See the papers listed in footnote 2. Detailed results for floating exchange rates are available from the author upon request.

¹⁹There is another somehow similar, but in two respects substantially different non-linear estimate in the literature by Koedijk *et al.* (1998). They derive the exchange rate function of a target zone currency in discrete time and arrive at the following estimated equation: $s_{t+1} - s_t = \beta_0 + \beta_1 x_t + \beta_2^- d^- x_t^2 + \beta_2^+ d^+ x_t^2 + \beta_3 x_t^3 + \varepsilon_{t+1}$, where d^- and d^+ are dummy variables, indicating whether the exchange rate is above or below the central parity. Hence, these authors discriminate the parameter of the quadratic term according to sign on the one hand, and employ a non-overlapping test equation on the other. We also carried out a Monte Carlo simulation on random walks for this particular specification and found almost no size distortions. For example, the sizes of the tests using 5 percent significance were almost equal to the nominal size: for β_1 , β_2^- , β_2^+ , and β_3 the size of the test was 5.0%, 4.9%, 6.2%, and 7.2%, respectively.

tests are larger, if the amount of overlap is larger, the autoregressive root of the data generating process is closer to unity, and the sample size is smaller. Of special importance, the biases in parameter estimation are in the direction of finding spurious mean reversion for unit root processes. Regarding the exchange rate application under study, we found that the biases are such to display all theoretical predictions of the Bertola and Svensson (1993) model for random walks and for linear stationary processes having roots close to unity. Moreover, hypothesis tests have substantial size distortions even when calculated with Newey-West HAC covariances. Consequently, statistical inference requires nonstandard distributions that we tabulated. Using critical values from these distributions, virtually none of the empirical results of the whole literature building on the drift-adjustment method are significant. Hence, our results are in the spirit of the findings of Ang and Bekaert (2005), who showed that previous favorable results on the predictability of long run stock returns were the consequence of small sample bias and size distortions.

The general conclusion is that much care should be taken when working with overlapping observations, because the magnitude of bias in parameter estimates and size distortions of hypothesis tests are generally unknown. Since all of these magnitudes depend on the amount of overlap, the sample size, and the autoregressive root of the data generating process, there is no unique set of critical values to be used for inference, but all applications should be accompanied with the analysis of the statistical properties of the particular model used.

References

- [1] Ang, A. and Bekaert, G. (2005). 'Stock return predictability: Is it there?' Revised version of NBER Working Paper No. 8207
- [2] Bertola, G. and Svensson, L.E.O. (1993). 'Stochastic devaluation risk and the empirical fit of target zone models', *Review of Economic Studies*, Vol. 60, pp. 689-712.

- [3] Britten-Jones, M. and Neuberger, A. (2004). *Improved inference and estimation in regression with overlapping observations*, EFA 2004 Maastricht Meetings Paper No. 4156.
- [4] Caramazza, F. (1993). 'French-German interest rate differentials and time-varying realignment risk', IMF Staff Papers, Vol. 40, pp.567-583.
- [5] Chen, Z. and Giovannini, A. (1997). 'The determinants of realignment expectations under the EMS: some empirical regularities', *European Economic Review*, Vol. 41, pp. 1687-1707.
- [6] Dickey, D.A. and Fuller, W.A. (1979). 'Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, Vol. 74, pp. 427-431.
- [7] Fama, E.F. and French, K.R. (1988). 'Permanent and temporary components in stock prices', *Journal of Political Economy*, Vol. 96, pp. 246-273.
- [8] Frankel, J.A. and Phillips, S. (1992). 'The European Monetary System: credible at last?' *Oxford Economic Papers*, Vol. 44, pp. 791-816.
- [9] Frankel, J.A. and Rose, A.K. (1995). 'Empirical research on nominal exchange rates', in: Grossman, G.M. and Rogoff, K. (eds.), *Handbook of International Economics vol. III*, Amsterdam: North Holland.
- [10] Garber, P.M. and Svensson, L.E.O. (1995). 'The operation and collapse of fixed exchange rate regimes', in: Grossman, G.M. and Rogoff, K. (eds.), *Handbook of International Economics vol. III*, Amsterdam: North Holland, pp. 1865-1911.
- [11] Hamilton, J.D. (1994). *Time series analysis*, Princeton: Princeton University Press.
- [12] Hansen, L.P. and Hodrick, R.J. (1980). 'Forward exchange rates as optimal predictors of future spot rates: An econometric analysis', *Journal of Political Economy*, Vol. 88, pp. 829-853.
- [13] Hodrick, R.J. (1992). 'Dividend yield and expected stock returns: alternative procedures for inference and measurement', *Review of Financial Studies*, Vol. 5, pp. 357-386.
- [14] Iannizzotto, M. and Taylor, M.P. (1999). 'The target zone model, non-linearity and mean

- reversion: Is the honeymoon really over?' *The Economic Journal*, Vol. 109, pp. C96-110.
- [15] Knot, K.H.W. and Sturm, J.E. (1999). 'Virtual and composite fundamentals in the ERM', *Scandinavian Journal of Economics*, Vol. 101, pp. 277-296.
- [16] Knot, K.H.W., Sturm, J.E. and de Haan, J. (1998). 'The credibility of the European exchange rate mechanism', *Oxford Economic Papers*, Vol. 50, pp. 186-200.
- [17] Koedijk, K., Stork P.A. and de Vries, C.G. (1998). 'An EMS target zone model in discrete time', *Journal of Applied Econometrics*, Vol. 13, pp. 31-48.
- [18] Kirby, C. (1997). 'Measuring the Predictable Variation in Stock and Bond Returns', *The Review of Financial Studies*, Vol. 10, pp. 579-630.
- [19] Krugman, P. (1991). 'Target zones and exchange rate dynamics', *Quarterly Journal of Economics*, Vol. 106, pp. 669-682.
- [20] Lindberg, H., Söderlind, P. and Svensson, L.E.O. (1993). 'Devaluation expectations: the Swedish krona 1985-92', *The Economic Journal*, Vol. 103, pp. 1170-1179.
- [21] Lundbergh, S. and Teräsvirta, T. (2006). 'A time series model for an exchange rate in a target zone with applications', *Journal of Econometrics* (forthcoming)
- [22] Mizrach, B. (1995). Target zone models with stochastic realignments: an econometric evaluation", *Journal of International Money and Finance*, Vol. 14, pp. 641-657.
- [23] Mundaca, G. (2004). 'Modelling probabilities of devaluations', *Economica*, Vol. 71, pp. 13-37.
- [24] Nelson, C.R. and Kim, M.J. (1993). 'Predictable stock returns: the role of small sample bias', *The Journal of Finance*, Vol. 48, pp. 641-661.
- [25] Phillips, P.C.B. (1987). 'Time series regression with a unit root', *Econometrica*, Vol. 55, pp. 277-301.
- [26] Phillips, P.C.B., McFarland, J.W. and McMahon, P.C. (1996). 'Robust tests of forward exchange market efficiency with empirical evidence from the 1920s', *Journal of Applied Econo-*

metrics, Vol. 11, pp. 1-22.

- [27] Richardson, M. and Smith, T. (1991). 'Tests of financial models in the presence of overlapping observations', *Review of Financial Studies*, Vol. 4, pp. 227-254.
- [28] Richardson, M. and Stock, J.H. (1989). 'Drawing in inferences from statistics based on multi-year asset returns', *Journal of Financial Economics*, Vol. 25, pp. 323-348.
- [29] Rose, A.K. and Svensson, L.E.O. (1994). 'European exchange rate credibility before the fall', *European Economic Review*, Vol. 38, pp. 1185-1216.
- [30] Rose, A.K. and Svensson, L.E.O. (1995). 'Expected and predicted realignments: The FF/DM exchange rate during the EMS', *Scandinavian Journal Economics*, pp. 173-200.
- [31] Smith, J. and Yadav, S. (1996). 'A comparison of alternative covariance matrices for models with over-lapping observations', *Journal of International Money and Finance*, Vol. 15, pp. 813-823.
- [32] Svensson, L.E.O. (1993). 'Assessing target zone credibility: Mean reversion and devaluation expectations in the ERM, 1979-1992', *European Economic Review*, Vol. 37, pp. 763-802.
- [33] Thomas, A.H. (1994). 'Expected devaluation and economic fundamentals', *IMF Staff Papers*, Vol. 41, pp. 262-285.
- [34] Werner, A.M. (1996). 'Target zones and realignment expectations: the Israeli and Mexican experiences', *IMF Staff Papers*, Vol. 43, pp. 571-586.

Table 1: The effects of the sample size

T	$\hat{\beta}_1^{(AR)}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$ \hat{\beta}_1^{(AR)} - 1 $	$ \hat{\beta}_1 - 1 $	$ \hat{\beta}_2 $	$ \hat{\beta}_3 $	$R^2_{(AR)}$	$R^2_{(PAR)}$
100	0.309	0.417	-0.535	-70.502	0.708	0.911	21.453	277.638	0.127	0.257
300	0.678	0.564	-0.059	-54.190	0.327	0.586	9.295	76.410	0.520	0.561
1000	0.889	0.774	-0.013	-11.840	0.113	0.277	2.730	13.034	0.828	0.838
3000	0.961	0.906	0.010	-1.866	0.039	0.112	0.677	1.960	0.940	0.941
10000	0.988	0.968	0.003	-0.194	0.012	0.037	0.126	0.200	0.982	0.982
30000	0.996	0.989	0.001	-0.023	0.004	0.013	0.025	0.023	0.994	0.994
100000	0.999	0.997	0.000	-0.002	0.001	0.004	0.004	0.002	0.998	0.998
T	$\hat{\beta}_1^{(AR)} < 1$	$\hat{\beta}_1 < 1$	$\hat{\beta}_2 < 0$	$\hat{\beta}_3 < 0$	$S(t_{\hat{\beta}_1^{(AR)}})$	$S(t_{\hat{\beta}_1})$	$S(t_{\hat{\beta}_2})$	$S(t_{\hat{\beta}_3})$	$S(W1)$	$S(W2)$
100	0.950	0.744	0.506	0.629	0.620	0.217	0.199	0.206	0.995	0.448
300	0.955	0.750	0.504	0.806	0.380	0.120	0.223	0.279	0.962	0.375
1000	0.956	0.755	0.497	0.901	0.244	0.090	0.279	0.363	0.842	0.367
3000	0.958	0.749	0.498	0.928	0.197	0.078	0.298	0.391	0.700	0.346
10000	0.958	0.743	0.495	0.941	0.170	0.060	0.295	0.389	0.599	0.324
30000	0.950	0.752	0.490	0.947	0.169	0.054	0.287	0.381	0.562	0.308
100000	0.962	0.750	0.500	0.941	0.162	0.053	0.287	0.375	0.551	0.298

Notes. Number of Monte Carlo draws: $N = 10000$ for each sample size T . Data generating process: $x_{t+1} = x_t + \varepsilon_{t+1}$, $\varepsilon_t \sim iidN(0, \sigma^2)$, $t = 1, 2, \dots, T + 22$, $x_1 = 0$, where $\sigma = 0.006821$, the standard deviation of daily logarithmic changes of the dollar/mark rate. Least squares estimation of $x_{t+22} = \beta_0^{(AR)} + \beta_1^{(AR)}x_t + u_{t+22}$ and $x_{t+22} = \beta_0 + \beta_1x_t + \beta_2x_t^2 + \beta_3x_t^3 + v_{t+22}$. $\hat{\beta}_i$: mean estimate for β_i ; $|\hat{\beta}_i - \beta_i|$: mean absolute bias. R^2 : median R^2 ; $\hat{\beta}_i < \beta_i$: empirical probability of less than true estimate. $S(t_{\hat{\beta}_i})$: size of the t -test of $H_0: \hat{\beta}_i - \beta_i = 0$ calculated with Newey-West HAC covariances (21 lags) at 5% nominal size, using the one sided Dickey-Fuller critical value of -2.86 for $\beta_1^{(AR)}$ and β_1 and the two-sided critical value of the t -distribution for β_2 and β_3 . $S(W1)$: size of the Wald-type χ^2 -test of $H_0: \beta_0 = \beta_1 - 1 = \beta_2 = \beta_3 = 0$ at 5% nominal size with Newey-West HAC covariances (21 lags). $S(W)$: size of the Wald-type χ^2 -test of $H_0: \beta_2 = \beta_3 = 0$ at 5% nominal size with Newey-West HAC covariances (21 lags).

Table 2: Summary of main findings

	shape of the distribution	bias	size of hypothesis tests
β_1^{AR}	skewed left; variance increases in k and λ , decreases in T ; $P(\text{est} < \text{true}) > 50\%$ and is constant in k and T , increases in λ	downward biased; bias increases in k and λ , decreases in T	size distortions; empirical size increases in k and λ , decreases in T
β_1	skewed left; variance increases in k and λ , decreases in T ; $P(\text{est} < \text{true}) > 50\%$ and is constant in k and T , increases in λ	downward biased when $\lambda = 1$, generally downward biased when $\lambda < 1$; bias increases in k and λ , decreases in T	size distortions; empirical size increases in k and λ , decreases in T
β_2	symmetric; variance increase in k and λ , decrease in T , $P(\text{est} < \text{true}) = 50\%$	no bias	size distortions; empirical size increases in k , λ , and T
β_3	skewed left; variance increases in k and λ , decreases in T ; $P(\text{est} < \text{true}) > 50\%$ and is constant in k , increases in T , not monotonic in λ	downward biased; bias increases in k when $\lambda = 1$, not monotonic in k when $\lambda < 1$, decreases in T , not monotonic in λ	size distortions; empirical size increases in k , λ , and T

Notes. Entries summarize the qualitative features of the results of our Monte Carlo simulations using the data generating process: $x_{t+1} = \lambda x_t + \varepsilon_{t+1}$, $\varepsilon_t \sim iidN(0, \sigma^2)$, $t = 1, 2, \dots, T + k$, $x_1 = 0$, $\sigma = 0.006821$ (which is the standard deviation of daily logarithmic changes of the dollar/mark rate), and least squares estimation of the overlapping linear autoregression, $x_{t+k} = \beta_0^{(AR)} + \beta_1^{(AR)} x_t + u_{t+k}$, and the overlapping polynomial autoregression, $x_{t+k} = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + v_{t+k}$. $P(\text{est} < \text{true})$ indicates the probability that the estimate is less than the true value. The true values are: $\beta_1^{(AR)} = \beta_1 = \lambda^k$, $\beta_2 = \beta_3 = 0$.

Figure 1: Theoretical predictions of the Bertola-Svensson model and their empirical fit for floating exchange rates

Notes. Source of the theoretical figures in Panels A-B-C: Bertola and Svensson (1993) p. 693-7, 708-9 and Rose and Svensson (1995) p.179, 181, 194, which were reproduced here using the baseline estimation results of Rose and Svensson (1995). The corresponding figures for the log-level of the dollar/mark rate and the Australian dollar/Swiss frank rate on Panels D to I are based on our estimates.

Figure 2: The effects of the amount of overlap on the distribution when there is a unit root

Notes: Data generating process: $x_t = x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$, $t = 2, \dots, 3000 + k$, $x_1 = 0$, where $\sigma = 0.006821$, the standard deviation of daily logarithmic changes of the dollar/mark rate. Least squares estimation of $x_{t+k} = \beta_0^{(AR)} + \beta_1^{(AR)} x_t + u_{t+k}$ and $x_{t+k} = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + v_{t+k}$. Number of Monte Carlo draws: 50000 for each k . **Panel A:** Kernel density estimate of the distribution of $\hat{\beta}_1^{(AR)}$. **Panel B:** Kernel density estimate of the distribution of $\hat{\beta}_1$. **Panel C:** Kernel density estimate of the distribution of $\hat{\beta}_2$. **Panel D:** Kernel density estimate of the distribution of $\hat{\beta}_3$.

Figure 3: The effects of the amount of overlap when there is a unit root

Notes: Data generating process: $x_t = x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$, $t = 2, \dots, 3000 + k$, $x_1 = 0$, where $\sigma = 0.006821$, the standard deviation of daily logarithmic changes of the dollar/mark rate. Least squares estimation of $x_{t+k} = \beta_0^{(AR)} + \beta_1^{(AR)} x_t + u_{t+k}$ and $x_{t+k} = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + v_{t+k}$. Number of Monte Carlo draws: 50000 for each k . **Panel A:** $P(\hat{\beta}_i < true)$: proportion of $\hat{\beta}_i$ estimates less than the population value. R^2 : mean R^2 of the polynomial autoregression. **Panel B:** $B(\hat{\beta}_i)$: mean bias of $\hat{\beta}_i$ estimates. **Panel C:** $S(t, \beta_i, 5\%)$: size of the t -test for β_i calculated with Newey-West HAC covariances (lags = $k - 1$) at 5% nominal size, using the one sided Dickey-Fuller critical value of -2.86 for $\beta_1^{(AR)}$ and β_1 and the two-sided critical value of the t -distribution for β_2

and β_3 . **Panel D:** $S(W1, 5\%)$: size of the Wald-type χ^2 -test for $H_0 : \beta_0 = \beta_1 - 1 = \beta_2 = \beta_3 = 0$, $S(W2, 5\%)$: size of the Wald-type χ^2 -test for $H_0 : \beta_2 = \beta_3 = 0$, both at 5% nominal size with Newey-West covariances (lags = $k - 1$).

Figure 4: The effects of the amount of overlap when the process is stationary

Notes: Data generating process: $x_t = 0.9x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$, $t = 2, \dots, 3000+k$, $x_1 = 0$, where $\sigma = 0.006821$, the standard deviation of daily logarithmic changes of the dollar/mark rate. Least squares estimation of $x_{t+k} = \beta_0^{(AR)} + \beta_1^{(AR)}x_t + u_{t+k}$ and $x_{t+k} = \beta_0 + \beta_1x_t + \beta_2x_t^2 + \beta_3x_t^3 + v_{t+k}$. Number of Monte Carlo draws: 50000 for each k . **Panel A:** $P(\hat{\beta}_i < true)$: proportion of $\hat{\beta}_i$ estimates less than the population value. R^2 : mean R^2 of the polynomial autoregression. **Panel B:** $B(\hat{\beta}_i)$: mean bias of $\hat{\beta}_i$ estimates. **Panel C:** $S(t, \beta_i, 5\%)$: size of the t -test for β_i calculated with Newey-West HAC covariances (lags = $k - 1$) at 5% nominal size, using the two-sided critical value of the t -distribution. **Panel D:** $S(W1, 5\%)$: size of the Wald-type χ^2 -test for $H_0 : \beta_0 =$

$\beta_1 - 0.9^k = \beta_2 = \beta_3 = 0$, $S(W2, 5\%)$: size of the Wald-type χ^2 -test for $H_0 : \beta_2 = \beta_3 = 0$, both at 5% nominal size with Newey-West covariances (lags = $k - 1$).

Figure 5: The effects of the autoregressive root

Notes: Data generating process: $x_t = \lambda x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$, $t = 2, \dots, 3023$, $x_1 = 0$, where $\sigma = 0.006821$, the standard deviation of daily logarithmic changes of the dollar/mark rate. Least squares estimation of $x_{t+k} = \beta_0^{(AR)} + \beta_1^{(AR)} x_t + u_{t+k}$ and $x_{t+k} = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + v_{t+k}$ for $k = 22$. Number of Monte Carlo draws: 50000 for each λ . **Panel A:** $P(\hat{\beta}_i < true)$: proportion of $\hat{\beta}_i$ estimates less than the population value. R^2 : mean R^2 of the polynomial autoregression. **Panel B:** $B(\hat{\beta}_i)$: mean bias of $\hat{\beta}_i$ estimates. **Panel C:** $S(t, \beta_i, 5\%)$: size of the t -test for β_i calculated with Newey-West HAC covariances (lags = $k - 1$) at 5% nominal size, using the one sided Dickey-Fuller critical value of -2.86 for $\beta_1^{(AR)}$ and β_1 when $\lambda = 1$ and the two-sided critical value of the t -distribution for β_2 and β_3 and for $\beta_1^{(AR)}$ and β_1 when $\lambda < 1$. **Panel D:** $S(W1, 5\%)$: size of the

Wald-type χ^2 -test for $H_0 : \beta_0 = \beta_1 - \lambda^k = \beta_2 = \beta_3 = 0$, $S(W2, 5\%)$: size of the Wald-type χ^2 -test for $H_0 : \beta_2 = \beta_3 = 0$, both at 5% nominal size with Newey-West HAC covariances (lags = $k - 1$).

DEPARTMENT OF MATHEMATICAL ECONOMICS AND ECONOMIC ANALYSIS

WORKING PAPERS IN ENGLISH

- 2007 / 1 Zsolt Darvas: *Estimation Bias and Inference in Overlapping Autoregressions: Implications for the Target Zone Literature*
- 2006 / 3 Zsolt Darvas – Zoltán Schepp: *Long Maturity Forward Rates of Major Currencies Are Stationary*
- 2005 / 5 Zsolt Darvas – Gábor Vadas: *A New Method For Combining Detrending Techniques with Application to Business Cycle Synchronization of the New EU Members*
- 2005 / 4 Zsolt Darvas – Andrew K. Rose – György Szapáry: *Fiscal Divergence and Business Cycle Synchronization: Irresponsibility is Idiosyncratic*
- 2004 / 1 József Móczár: *Goodwin Cycles, Bifurcations and Complex Dynamics in the Nonlinear Domar-Harrod-Leontief Systems*
- 2003 / 5 Ernő Zalai: *The von Neumann Model and the Early Models of General Equilibrium*
- 2003 / 4 József Móczár: *Harrod Model is Structurally Stable*